

Univerzita Karlova v Praze

Přírodovědecká fakulta

BAKALÁŘSKÁ PRÁCE



Jakub Kocák

Kvantová variační Monte Carlo metoda

Katedra fyzikální a makromolekulární chemie

Vedoucí bakalářské práce: RNDr. Filip Uhlík, Ph.D.

Studijní program: Chemie

Studijní obor: Chemie v přírodních vědách

Praha 2014

Podakovanie

Na tomto mieste by som veľmi rád poďakoval RNDr. Filipovi Uhlíkovi, Ph.D. za hodnotné stretnutia, obrovské množstvo trpezlivosti, veľkú ochotu a hojný počet e-mailov, ktoré mi pomohli k napísaniu tejto práce. Rovnako by som veľmi rád poďakoval priateľom a v neposlednom rade celej mojej rodine, ktorá ma celý čas podporovala.

Prohlášení:

Prohlašuji, že jsem závěrečnou práci zpracoval samostatně a že jsem uvedl všechny použité informační zdroje a literaturu. Tato práce ani její podstatná část nebyla předložena k získání jiného nebo stejného akademického titulu.

V Praze dne

Podpis

Název práce: Kvantová variační Monte Carlo metoda

Autor: Jakub Kocák

Katedra: Katedra fyzikální a makromolekulární chemie

Vedoucí bakalářské práce: RNDr. Filip Uhlík, Ph.D.

Abstrakt: Práce se zabývá studiem variační Monte Carlo metody ve kvantově-mechanických systémech. Udělali jsme rozbor volby tvaru zkusmé vlnové funkce a následně jsme zvolený tvar optimalizovali pro atom helia v singletovém a tripletovém stavu. V první kapitole jsou diskutovány základní pojmy kvantové mechaniky a obecné vlastnosti vlnové funkce i vlastnosti základního stavu systému. V druhé kapitole rozebíráme výpočetní algoritmy použité k vyčíslení integrálů, odhadu chyb a optimalizaci. V třetí kapitole prezentujeme výsledky optimalizace a vlastnosti zkusmé vlnové funkce s optimalizovanými parametry.

Klíčová slova: variační princip, Monte Carlo metoda, základní stav, helium

Title: Quantum Variational Monte Carlo method

Author: Jakub Kocák

Department: Department of Physical and Macromolecular Chemistry

Supervisor: RNDr. Filip Uhlík, Ph.D.

Abstract:

In this thesis, we study variational Monte Carlo method in quantum-mechanical systems. We analysed choice of trial wave function and afterwards we optimized selected function for helium singlet and triplet state. In the first chapter, we discuss basic notions of quantum mechanics and general properties of wave function and properties of ground state of the system. In the second chapter, we consider computational algorithms used to calculate integrals, to estimate errors and for optimization. In the third chapter, we present results of optimization and properties of optimized trial wave function.

Keywords: variational principle, Monte Carlo method, ground state, helium

Obsah

Úvod	2
1 Základný stav systému	3
1.1 Riešenie časovej a nečasovej Schrödingerovej rovnice	3
1.2 Vlastnosti vlnovej funkcie	5
1.2.1 Všeobecné vlastnosti vlnovej funkcie	5
1.2.2 Vlastné stavy	6
1.2.3 Nerozlíšiteľnosť častíc	8
1.2.4 Reálny Hamiltonián	9
1.2.5 Variačný princíp	10
1.2.6 Viriálový teorém	12
1.2.7 Asymptotické správanie v Coulombovskom potenciáli	14
1.3 Bornova-Oppenheimerova aproximácia	16
2 Výpočtové algoritmy	18
2.1 Monte Carlo metóda	18
2.1.1 Metropolisov-Hastingsov algoritmus	19
2.2 Závislosť vybraných vzoriek	20
2.2.1 Metóda blokovania	20
2.3 Optimalizačné metódy	21
2.3.1 Optimalizácia v jednorozmernom priestore	21
2.3.2 Optimalizácia vo viacrozmernom priestore	26
3 Výsledky, Systém He	30
3.1 Optimalizácia	30
3.1.1 Voľba tvaru vlnovej funkcie	33
3.1.2 Výber generovania vzoriek	33
3.1.3 Vlnová funkcia s jedným parametrom	37
3.1.4 Vlnová funkcia s viacerými parametrami	39
3.1.5 Vlastnosti optimalizovanej funkcie	48
Záver	59
4 Dodatky	60
4.1 Separácia premenných	60
4.2 Ortogonalita vlastných funkcií	60
4.3 Komutátory	61
4.4 Potenciál vo viacčasticovom systéme	64
4.5 Štatistika	66
Literatúra	70

Úvod

Od čias formulácie kvantovej mechaniky sa počet analytických riešení Schrödingerovej rovnice ukázal obmedzený. Preto v tomto poli sa ukázala nutná účasť výpočtovej techniky na približné riešenie kvantovo-mechanických systémov. V prípade systémoch atómov a molekúl nie je hľadanie optimálnej metódy, ktorá by s danou presnosťou a v rozumnom výpočtovom čase riešenie našla, dostatočne uspokojené. Veľkým problémom je rýchlo narastajúca zložitosť vlnovej funkcie s počtom častíc (elektrónov a jadier). Jednou z možností ponúka variačný princíp, ktorý ponúka možnosť výpočtu horného odhadu energie základného stavu systému. Obmedzenie sa na lineárnu kombináciu vlnových funkcií atómových orbitálov, pri ktorých sa integrály rátať jednoduchšie, si vyžaduje veľkú bázu. Výsledky však nie sú dostatočne uspokojivé.

Variačná metóda si však nedáva žiadne obmedzenie na skusmú vlnovú funkciu, a tým pádom nič nebráni zvoliť si vlastnú skusmú vlnovú funkciu s parametrami. To si však vyžaduje rátať všeobecne neanalytické integrály. Tu prichádza vhod rátanie integrálov Monte Carlo metódou, ktorej neprekáža ani zložitosť funkcie, ani veľký počet častíc. Cieľom tejto práce je vyskúšať si kvantovú variačnú Monte Carlo metódu na najjednoduchšom analyticky neriešiteľnom atóme – héliu.

Pre úplnosť uvádzam používanú notáciu:

- Polohové vektory častíc sú označené tučne \mathbf{r}_j a číslované sú dolným indexom j , aby sa rozlíšil od komplexnej jednotky i .
- Kvantovomechanické stavy sú podľa potreby označené buď konkrétne v súradnicovej reprezentácii vlnovou funkciou $\psi(\mathbf{r}_j, t)$ a k nej komplexne združenou funkciou $\psi^*(\mathbf{r}_j, t)$, alebo vo všeobecnosti (bez uvedenia reprezentácie) v Diracovej bra-ket notácii ($|\psi\rangle$ je "ket"-stav a k nemu združený $\langle\psi|$ je "bra"-stav).
- Ak máme N časticový systém, tak pre jednoduchosť vo vlnových funkciách namiesto označenia $\psi(\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_N)$ používame značenie $\psi(\mathbf{r}_j)$.

1. Základný stav systému

1.1 Riešenie časovej a nečasovej Schrödingerovej rovnice

V prípade riešenia vývoja systému na úrovni kvantovej mechaniky zaberá ústrednú rolu *časová Schrödingerova rovnica* [1]

$$i\hbar \frac{\partial}{\partial t} \Psi(\mathbf{r}_j, t) = \widehat{H}(t) \Psi(\mathbf{r}_j, t), \quad (1.1)$$

kde \hbar je redukovaná Planckova konštanta, $\widehat{H}(t)$ operátor Hamiltoniánu systému vo všeobecnosti závislý od času t a $\Psi(\mathbf{r}_j, t)$ vlnová funkcia systému.

V prípade, že Hamiltonián systému nezávisí od času, upravíme rovnicu na tvar

$$\left(i\hbar \frac{\partial}{\partial t} - \widehat{H} \right) \Psi(\mathbf{r}_j, t) = 0, \quad (1.2)$$

a riešime vlastný problém s vlastným číslom $\lambda = 0$. Podľa 4.1 môžeme pomocou separácie premenných $X = (t)$, $Y = (\mathbf{r}_j)$ rozdeliť problém na dva vlastné problémy s operátormi $\widehat{A} = i\hbar \frac{\partial}{\partial t}$ a $\widehat{B} = \widehat{H}$ ¹

$$i\hbar \frac{\partial}{\partial t} \varphi(t) = \lambda_A \varphi(t), \quad (1.3)$$

$$\widehat{H} \psi(\mathbf{r}_j) = E_n \psi_n(\mathbf{r}_j). \quad (1.4)$$

Výsledná vlnová funkcia $\Psi(\mathbf{r}_j, t)$ má potom pre vlastné stavy tvar

$$\Psi(\mathbf{r}_j, t) = \psi_n(\mathbf{r}_j) \varphi(t). \quad (1.5)$$

Keďže vlastné číslo rovnice (1.2) má byť nulové, dostávame určené vlastné číslo λ_A

$$\lambda_A = E_n. \quad (1.6)$$

Riešením diferenciálnej rovnice (1.3) dostávame riešenie

$$\varphi(t) = N e^{\frac{E_n t}{i\hbar}}, \quad (1.7)$$

kde N je normalizačná konštanta zväčša volená $N = 1$ za predpokladu, že funkcia $\psi_n(\mathbf{r}_j)$ už je normovaná.

Takže pre vlastné stavy riešenie časovej Schrödingerovej rovnice má tvar

$$\Psi(\mathbf{r}_j, t) = \psi_n(\mathbf{r}_j) e^{\frac{E_n t}{i\hbar}}. \quad (1.8)$$

Overíme ešte predpoklady. Oba operátory $i\hbar \frac{\partial}{\partial t}$ a \widehat{H} sú lineárne a keďže sme predpokladali, že Hamiltonián je nečasový, každý z operátorov pôsobí podľa disjunktných množín premenných ($i\hbar \frac{\partial}{\partial t}$ podľa času (t) , Hamiltonián podľa (\mathbf{r}_j)).

¹V prípade rovnice (1.3) môžeme parciálnu deriváciu nahradiť úplnou deriváciou, keďže funkcia $\varphi(t)$ je funkcia iba jednej premennej.

Našli sme síce iba riešenie vlastných (stacionárnych) stavov, no môžeme využiť výsledok 4.2. Predpokladajme, že množina vlastných funkcií $\psi_n(\mathbf{r}_j)$ Hamiltoniánu tvorí *úplnú ortonormálnu bázu daného Hilbertovho priestoru*. Teda, že spĺňajú

$$\langle \psi_m | \psi_n \rangle = \delta_{mn}, \quad (1.9)$$

kde δ_{mn} je Kroneckerovo delta.

Ak vlastné čísla (energie) Hamiltoniánu boli nedegenerované

$$(\forall m, n \in \mathbb{N})(E_m \neq E_n), \quad (1.10)$$

tak podľa 4.2 máme priamo ortogonálnu množinu funkcií. Ak sú však vlastné stavy Hamiltoniánu degenerované (čo sa bežne pri riešení stáva), tak v danom systéme vezmeme maximálnu množinu navzájom komutujúcich operátorov (vrátane Hamiltoniánu)² ($\hat{H}, \hat{A}, \hat{B}, \dots$). Vtedy majú tieto veličiny spoločnú množinu vlastných funkcií. Každý vlastnej funkcii pripadá množina vlastných čísel (E_m, A_m, B_m, \dots). Keďže merateľné veličiny majú reálne vlastné čísla (hermitovské operátory) a budeme porovnávať vlastné čísla dvoch vlastných funkcií, tak rozdielnosť aspoň v jednom vlastnom čísle implikuje podľa 4.2 ortogonalitu všetkých vlastných funkcií aj v prípade degenerácie

$$(\forall m, n \in \mathbb{N})((E_m, A_m, B_m, \dots) \neq (E_n, A_n, B_n, \dots)). \quad (1.11)$$

Ortonormalitu zabezpečíme normovaním vlastných funkcií. Ak máme vlastnú funkciu \tilde{f}_m , jej normovaním dostaneme vlastnú funkciu f_m s normou 1

$$f_m = \frac{\hat{f}_m}{\langle \hat{f}_m | \hat{f}_m \rangle^{1/2}}, \quad (1.12)$$

$$\langle f_m | f_m \rangle = \frac{\langle \hat{f}_m | \hat{f}_m \rangle}{\langle \hat{f}_m | \hat{f}_m \rangle^{1/2} \langle \hat{f}_m | \hat{f}_m \rangle^{1/2}} = 1. \quad (1.13)$$

Keď teda máme úplnú ortonormálnu bázu Hilbertovho priestoru $\{\psi_n\}$, môžeme hocikáku funkciu ψ z tohto priestoru vyjadriť ako lineárnu kombináciu v tejto báze³

$$|\psi\rangle = \sum_{n=1}^{\infty} c_n |\psi_n\rangle. \quad (1.14)$$

kde c_n sú komplexné koeficienty. Výpočtom skalárneho súčinu $\langle \psi_m | \psi \rangle$ dostaneme pre danú funkciu ψ koeficienty

$$\langle \psi_m | \psi \rangle = \left\langle \psi_m \left| \sum_{n=1}^{\infty} c_n \psi_n \right. \right\rangle = \sum_{n=1}^{\infty} c_n \delta_{mn} = c_m. \quad (1.15)$$

Takže ak máme v nejakom čase t_0 danú stav kvantovomechanického systému, rozložíme ho do úplnej ortonormálnej báze Hilbertovho priestoru vlastných stavov

²Nazýva sa aj *úplná množina pozorovateľných veličín*.

³Budeme uvádzať vzťahy v prípade diskretných stavov, v prípade spojitého spektra platia analogické integrálne tvary vzťahov.

Hamiltoniánu, ktorých časový vývoj už poznáme. Potom ak máme danú funkciu $\Psi(\mathbf{r}_j, t_0)$, riešenie časovej Schrödingerovej rovnice má tvar

$$|\Psi(\mathbf{r}_j, t)\rangle = \sum_{n=1}^{\infty} |\psi_n(\mathbf{r}_j)\rangle \langle\psi_n(\mathbf{r}_j) | \Psi(\mathbf{r}_j, t_0)\rangle e^{\frac{E_n t}{i\hbar}}, \quad (1.16)$$

pričom sme využili, že časová Schrödingerová rovnica je lineárna rovnica a rovnako lineárne sa správajú aj jej riešenia.

Zúžili sme teda problém riešenia časovej Schrödingerovej rovnice na problém riešenia nečasovej Schrödingerovej rovnice, ktorý hľadá vlastné stavy (vlastný problém) daného systému

$$\widehat{H} |\psi_n(\mathbf{r}_j)\rangle = E_n |\psi_n(\mathbf{r}_j)\rangle. \quad (1.17)$$

Ak však máme systém N častíc, vlnová funkcia je funkcia

$$\psi : \mathbb{R}^{3N} \rightarrow \mathbb{C}. \quad (1.18)$$

Rapidne zvyšovanie dimenzie konfiguračného priestoru systému s počtom častíc je hlavným dôvodom komplikácie riešenia vlastného problému.

1.2 Vlastnosti vlnovej funkcie

Pri variačnej metóde sa vlnová funkcia $\psi(\mathbf{r}_j)$ volí pred samotným výpočtom, preto je potrebné zistiť, aké základné vlastnosti musí vlnová funkcia spĺňať.

1.2.1 Všeobecné vlastnosti vlnovej funkcie

V prvom rade musí to byť komplexná funkcia $\psi(\mathbf{r}_j)$ definovaná na celom konfiguračnom priestore \mathbb{R}^{3N} [1]

$$\psi : \mathbb{R}^{3N} \rightarrow \mathbb{C}^*. \quad (1.19)$$

Teda musí byť v každom bode konfiguračného priestoru daná práve jedna hodnota⁴. Ak je v danom mieste vlnová funkcia nekonečná, musíme mať na pamäti *pravdepodobnostnú interpretáciu* vlnovej funkcie, t.j. pravdepodobnosť $dp(\mathbf{r}_j)$ nájdenia častice/častíc v objemovom elemente konfiguračného priestoru dV v bode $((\mathbf{r}_j))$ je rovná

$$dp(\mathbf{r}_j) = |\psi(\mathbf{r}_j)|^2 dV. \quad (1.21)$$

Ak by teda napríklad vlnová funkcia nadobudla nekonečnú hodnotu iba v jednom bode, tak nakoniec je príspevok tohto bodu nulový, keďže jeden bod je množina miery nula, a teda pri integrovaní hodnotu integrálu nezmení⁵.

⁴Množina \mathbb{C}^* je množina komplexných čísel \mathbb{C} rozšírená o nekonečno ∞ a o algebru s nekonečnom

$$\mathbb{C}^* := \mathbb{C} \cup \{\infty\}. \quad (1.20)$$

⁵Problém je samozrejme komplikovanejší a rozoberieme ho ďalej, keďže vďaka ďalším vlastnostiam vlnovej funkcie nemusí byť integrál cez takúto divergentnú oblasť nulový, ani konečný.

Hlavnou vlastnosťou prichádzajúcou s platnosťou pravdepodobnostnej interpretácie je kvadratická integrabilita vlnovej funkcie. Táto vlastnosť súvisí s tým, že keď vypočítame pravdepodobnosť výskytu v celom konfiguračnom priestore $V = \mathbb{R}^{3N}$, tak musíme dostať jednotku - celková pravdepodobnosť, že systém v nejakom konfiguračnom stave nájdeme, je jedna

$$p(V) = \int_V dp(\mathbf{r}_j) = \int_V |\psi(\mathbf{r}_j)|^2 dV = 1. \quad (1.22)$$

Z rovnice vidíme, že vlnová funkcia $\psi(\mathbf{r}_j)$ musí pri integrovaní v kvadráte veľkosti cez celý priestor V dávať konečnú hodnotu (kvadraticky integrabilná). Ak by sme dostali konštantu inú ako 1, tak riešenie je jednoduché. Rovnako, ako v prípade (1.13), by sme starú vlnovú funkciu $\psi(\mathbf{r}_j)$ znormovali.

1.2.2 Vlastné stavy

Nás budú, samozrejme, zaujímať vlnové funkcie, ktoré sú riešením nečasovej Schrödingerovej rovnice (1.13), vlastné vlnové funkcie. V tomto prípade musí vlnová funkcia spĺňať nasledujúce tri vlastnosti. Musí byť *konečná*, *spojitá* a *spojitá v parciálnych deriváciách* $\frac{\partial\psi(\mathbf{r}_j)}{\partial x_j}$, $\frac{\partial\psi(\mathbf{r}_j)}{\partial y_j}$ a $\frac{\partial\psi(\mathbf{r}_j)}{\partial z_j}$ v prípade konečného potenciálu.

Posledná vlastnosť súvisí s riešením nečasovej Schrödingerovej rovnice a hustotou toku pravdepodobnosti. *Hustota toku pravdepodobnosti* \mathbf{j}_k jednej k -tej častice [1] je definovaná ako

$$\mathbf{j}_k := \frac{\hbar}{2m_k i} (\psi^*(\mathbf{r}_j) \nabla_k \psi(\mathbf{r}_j) - \psi(\mathbf{r}_j) \nabla_k \psi^*(\mathbf{r}_j)). \quad (1.23)$$

Na hustotu toku pravdepodobnosti kladieme podmienku spojitosti, lenže tá je splnená, ak sú spojité aj funkcie v jej definícii. V gradiente vystupujú parciálne derivácie podľa súradníc, preto aj tie musia byť spojité.

Druhým odôvodnením, ktoré zároveň aj objasňuje podmienku konečných zmien potenciálu, je nasledujúci rozbor jednorozmerného prípadu. Predpokladajme jednorozmerný prípad, teda vlnovú funkciu $\psi(x)$ a nerelativistický Hamiltonián vo tvare⁶

$$\widehat{H} = -\frac{\hbar^2}{2m} \frac{\partial^2}{\partial x^2} + V(x), \quad (1.24)$$

kde m je hmotnosť častice a $V(x)$ je potenciál častice.

V prípade nečasovej Schrödingerovej rovnice dostaneme rovnicu

$$-\frac{\hbar^2}{2m} \frac{\partial^2 \psi}{\partial x^2}(x) + V(x)\psi(x) = E\psi(x). \quad (1.25)$$

Upravíme ju na tvar

$$\frac{\partial^2 \psi}{\partial x^2}(x) = \frac{2m}{\hbar^2} (V(x) - E) \psi(x). \quad (1.26)$$

⁶Kedže ide o jednorozmerný prípad, parciálne derivácie sa rovnajú úplným deriváciám, ale pre väčšiu názornosť sme ponechali parciálne derivácie.

Vyberieme si bod $x = a$ a jeho malé ε -okolie $(a - \varepsilon, a + \varepsilon)$ ⁷, na ktorom danú rovnicu zintegrujeme

$$\int_{a-\varepsilon}^{a+\varepsilon} \frac{\partial^2 \psi}{\partial x^2}(x) dx = \int_{a-\varepsilon}^{a+\varepsilon} \frac{2m}{\hbar^2} (V(x) - E) \psi(x) dx,$$

$$\left| \frac{\partial \psi}{\partial x}(a + \varepsilon) - \frac{\partial \psi}{\partial x}(a - \varepsilon) \right| = \frac{2m}{\hbar^2} \left| \int_{a-\varepsilon}^{a+\varepsilon} (V(x) - E) \psi(x) dx \right|,$$

kde sme v poslednom kroku urobili absolútnu hodnotu z oboch strán rovnice.

Teraz použijeme predpoklad konečného potenciálu. Ak je na intervale $(a - \varepsilon, a + \varepsilon)$ potenciál konečný⁸ a rovnako je konečná vlnová funkcia, tak potom môžeme definovať K ako maximum integrovaného výrazu

$$K := \max\{|(V(x) - E) \psi(x)| \mid x \in (a - \varepsilon, a + \varepsilon)\} < \infty \quad (1.27)$$

Môžeme potom písať

$$\left| \int_{a-\varepsilon}^{a+\varepsilon} (V(x) - E) \psi(x) dx \right| \leq \int_{a-\varepsilon}^{a+\varepsilon} |(V(x) - E) \psi(x)| dx \leq \int_{a-\varepsilon}^{a+\varepsilon} K dx = 2K\varepsilon. \quad (1.28)$$

Ak použijeme limitu $\varepsilon \rightarrow 0$, tak dostaneme

$$\lim_{\varepsilon \rightarrow 0} \left| \frac{\partial \psi}{\partial x}(a + \varepsilon) - \frac{\partial \psi}{\partial x}(a - \varepsilon) \right| = 0, \quad (1.29)$$

keďže výraz je menší ako $2K\varepsilon$, a teda v limite ide do nuly. Preto aj limita výrazu v absolútnej hodnote je 0

$$\lim_{\varepsilon \rightarrow 0} \left(\frac{\partial \psi}{\partial x}(a + \varepsilon) - \frac{\partial \psi}{\partial x}(a - \varepsilon) \right) = 0. \quad (1.30)$$

Za predpokladu konečnosti jednotlivých limít dostávame podľa vety o aritmetike limít

$$\lim_{\varepsilon \rightarrow 0} \frac{\partial \psi}{\partial x}(a + \varepsilon) = \lim_{\varepsilon \rightarrow 0} \frac{\partial \psi}{\partial x}(a - \varepsilon), \quad (1.31)$$

teda parciálne derivácie sa rovnajú pri limite sprava i zľava. Preto sú parciálne derivácie v bode a spojité. Bod nebol špecificky vybraný, preto to platí pre všetky x .

Uvedený príklad v jednorozmernom prípade ukazuje, ako predpoklad konečného potenciálu vplyva na spojitosť parciálnych derivácií.

Keď už má funkcia spojité parciálne derivácie, tak už samotná funkcia musí byť spojitá.

Ďalšou vlastnosťou, ktorou sa vyznačujú vlastné stavy, je stredná hodnota operátora $\widehat{H} - E_n$, resp. $(\widehat{H} - E_n)^l$, kde $l \in \mathbb{N}$ je prirodzené číslo. Keď budeme vychádzať z rovnice (1.4), tak jednoducho na základe linearity dostaneme

$$(\widehat{H} - E_n) \psi_n(\mathbf{r}_j) = 0, \quad (1.32)$$

⁷ $\varepsilon > 0$.

⁸Odcítanie konečnej konštanty E nezmení konečnosť výrazu.

vo význame nulovej funkcie. Túto rovnosť pre všetky vyššie mocniny dostaneme použitým tejto rovnosti pri pôsobení prvého člena $(\widehat{H} - E_n)$

$$(\widehat{H} - E_n)^l \psi_n(\mathbf{r}_j) = (\widehat{H} - E_n)^{l-1} (\widehat{H} - E_n) \psi_n(\mathbf{r}_j) = (\widehat{H} - E_n)^{l-1} 0(\mathbf{r}_j) = 0. \quad (1.33)$$

Pri tejto vlastnosti je potrebné poznať hodnotu E_n , ktorú pri praktickom výpočte hľadáme. Preto ako túto hodnotu volíme E , teda strednú hodnotu energie našej funkcie. Pri tejto voľbe je pre $l = 1$ rovnosť strednej hodnoty nule zabezpečená stále

$$\langle \psi | (\widehat{H} - E) \psi \rangle = \langle \psi | \widehat{H} \psi \rangle - E \langle \psi | \psi \rangle = 0. \quad (1.34)$$

Pre $l = 2$ nás bude preto zaujímať hodnota

$$\Delta H := \sqrt{\langle \psi | (\widehat{H} - E)^2 \psi \rangle}. \quad (1.35)$$

Pre $l = 2$ môžeme využiť, že operátor $(\widehat{H} - E_n)$ je hermitovský⁹. Potom dostaneme

$$\langle \psi | (\widehat{H} - E_n)^2 \psi \rangle = \langle (\widehat{H} - E_n) \psi | (\widehat{H} - E_n) \psi \rangle. \quad (1.36)$$

1.2.3 Nerozlíšiteľnosť častíc

Dôležitou vlastnosťou vlnovej funkcie je jej správanie pri zámene identických častíc. Pred rozborom tohto problému však budeme musieť urobiť ešte dve veci, explicitne vypísať v premenných vlnovej funkcie aj spinové premenné a definovať operátor zámenny (permutácie) častíc. Polohové premenné j -tej častice sú označené \mathbf{r}_j , spinové premenné ako ξ_j . Vlnová funkcia potom vyzerá ako $\psi(\mathbf{r}_j, \xi_j)$.

Operátor zámenny \widehat{P}_{kl} častíc k a l je definovaný takto

$$\widehat{P}_{kl} \psi(\dots, \mathbf{r}_k, \xi_k, \dots, \mathbf{r}_l, \xi_l, \dots) := \psi(\dots, \mathbf{r}_l, \xi_l, \dots, \mathbf{r}_k, \xi_k, \dots). \quad (1.37)$$

Vo vlnovej funkcii vystupujú premenné jednotlivých častíc v usporiadanom poradí. Avšak podľa princípu nerozlíšiteľnosti častíc sa pri zámene dvoch nesmie zmeniť fyzikálne merateľné vlastnosti systému. Aby teda nová vlnová funkcia spĺňala rovnaké rovnice (a napríklad, aby mala rovnakú funkciu hustoty pravdepodobnosti), môže byť vlnová funkcia po zámene častíc komplexným K -násobkom pôvodnej vlnovej funkcie

$$\widehat{P}_{kl} \psi(\mathbf{r}_j, \xi_j) := K \psi(\mathbf{r}_j, \xi_j). \quad (1.38)$$

Jeho možné hodnoty sa však dajú jednoducho určiť. Vieme, že pri opätovnom pôsobení operátora zámenny častíc musíme dostať pôvodnú funkciu

$$\widehat{P}_{kl} \widehat{P}_{kl} \psi(\mathbf{r}_j, \xi_j) = \widehat{P}_{kl} K \psi(\mathbf{r}_j, \xi_j) = K^2 \psi(\mathbf{r}_j, \xi_j) = \psi(\mathbf{r}_j, \xi_j). \quad (1.39)$$

Je teda jasné, že jediné riešenia sú hodnoty $K \in \{+1, -1\}$. Z kvantovej teórie poľa potom vychádza [2], že ak ide o *bozóny*¹⁰, tak $K = +1$ (vlnové funkcie sa

⁹Operátor \widehat{H} je hermitovský, rovnako aj konštanta E_n je hermitovský operátor. Potom stačí už iba využiť, že lineárna kombinácia hermitovských operátorov je hermitovský operátor.

¹⁰Častice s celočíselným spinom.

nazývajú *symetrické* voči permutácii častíc) a ak ide o *fermióny*¹¹, tak $K = -1$ (vlnové funkcie sa nazývajú *antisymetrické* voči permutácii častíc).

Keďže sa budeme zaoberať vlnovými funkciami elektrónov, ktoré patria medzi fermióny, tak naša vlnová funkcia musí byť antisymetrická.

1.2.4 Reálny Hamiltonián

Niektoré vlastnosti nám nedávajú len reštrikcie na vlnovú funkciu, ale aj istú slobodu vo voľbe. V tomto odstavci ukážeme, že nemusíme uvažovať vo všeobecnosti komplexnú vlnovú funkciu, ale postačí nám reálna funkcia pri hľadaní vlastných stavov.

Ako už vieme, vlnová funkcia $\psi(\mathbf{r}_j)$ je vo všeobecnosti komplexná funkcia. Nás však zaujímajú vlastné stavy, teda riešenia nečasovej Schrödingerovej rovnice

$$\widehat{H}\psi_n(\mathbf{r}_j) = E_n\psi_n(\mathbf{r}_j). \quad (1.40)$$

Používaný Hamiltonián je nerelativistický s kinetickým a potenciálnym členom s N časticami

$$\widehat{H} := - \sum_{k=1}^N \frac{\hbar^2}{2m_k} \Delta_k + V(\mathbf{r}_j). \quad (1.41)$$

Keď vezmeme urobíme komplexné združenie oboch strán rovnice (1.40), tak dostaneme

$$\begin{aligned} (\widehat{H}\psi_n(\mathbf{r}_j))^* &= (E_n\psi_n(\mathbf{r}_j))^* , \\ \widehat{H}^*\psi_n^*(\mathbf{r}_j) &= E_n^*\psi_n^*(\mathbf{r}_j) , \\ \widehat{H}\psi_n^*(\mathbf{r}_j) &= E_n\psi_n^*(\mathbf{r}_j) , \end{aligned} \quad (1.42)$$

kde sme najprv využili distributivitu komplexného združenia, že vlastné čísla Hamiltoniánu (hermitovského operátora) sú reálne a že náš zvolený Hamiltonián je reálny. Keď však vezmeme rovnice (1.40) a (1.42) a sčítame ich, dostaneme

$$\widehat{H}(\psi_n(\mathbf{r}_j) + \psi_n^*(\mathbf{r}_j)) = E_n(\psi_n(\mathbf{r}_j) + \psi_n^*(\mathbf{r}_j)) , \quad (1.43)$$

kde sme využili linearitu Hamiltoniánu.

Teraz je jasné, že ak máme vlastnú funkciu $\psi_n(\mathbf{r}_j)$ Hamiltoniánu \widehat{H} , tak vieme z tejto funkcie vyrobiť funkciu $\tilde{\psi}_n(\mathbf{r}_j) := (\psi_n(\mathbf{r}_j) + \psi_n^*(\mathbf{r}_j))$, ktorá je takisto riešením (1.40) s rovnakým vlastným číslom a je reálna

$$\begin{aligned} \tilde{\psi}_n(\mathbf{r}_j)^* &= (\psi_n(\mathbf{r}_j) + \psi_n^*(\mathbf{r}_j))^* = \\ &= (\psi_n^*(\mathbf{r}_j) + \psi_n(\mathbf{r}_j)) = \\ &= \tilde{\psi}_n(\mathbf{r}_j) . \end{aligned}$$

Funkciu $\tilde{\psi}_n(\mathbf{r}_j)$ však treba ešte podľa (1.13) znormovať. To však nezmení, že je reálna, ani to, že rieši rovnicu (1.40).

Z toho potom vyplýva, že našu vlnovú funkciu si môžeme dovoliť voliť iba v obore reálnych vlnových funkcií.

¹¹Častice s poločíselným spinom.

1.2.5 Variačný princíp

V nasledujúcom odstavci si ukážeme variačný princíp v kvantovej mechanike, ktorý je jedným z kľúčových vzťahov využitých pri riešení nečasovej Schrödingerovej rovnice¹².

Predpokladajme, že máme daný systém, ktorý má úplnú ortonormálnu bázu vlastných funkcií $|\psi_n\rangle$ ¹³, ktorá je navyše usporiadaná podľa vlastných čísel, čo znamená

$$(\forall i, j \in \mathbb{N}_0) ((i > j) \Rightarrow (E_i \geq E_j)) . \quad (1.44)$$

Predpokladajme, že máme normovanú vlnovú funkciu daného systému $|\psi\rangle$, nie nutne vlastnú funkciu. Podľa (1.14) ju môžeme jednoznačne rozložiť do úplnej ortonormálnej báze vlastných funkcií

$$|\psi\rangle = \sum_{n=0}^{\infty} c_n |\psi_n\rangle . \quad (1.45)$$

Bude nás zaujímať stredná hodnota energie vlnovej funkcie $|\psi\rangle$. Podľa definície dostaneme

$$\langle \widehat{H} \rangle := \langle \psi | \widehat{H} \psi \rangle = \left\langle \sum_{n=0}^{\infty} c_n \psi_n \left| \widehat{H} \sum_{m=0}^{\infty} c_m \psi_m \right. \right\rangle . \quad (1.46)$$

Z linearity Hamiltoniánu, linearity skalárneho súčinu v druhom argumente, antilinearite skalárneho súčinu v prvom argumente a použitím toho, že funkcie $|\psi_m\rangle$ sú vlastné funkcie Hamiltoniánu, tak dostaneme

$$\langle \widehat{H} \rangle = \sum_{n=0}^{\infty} \sum_{m=0}^{\infty} c_n^* c_m E_m \langle \psi_n | \psi_m \rangle . \quad (1.47)$$

Funkcie ψ_n a ψ_m sú ortonormálne, preto vo výraze dostaneme Kroneckerovo delta

$$\begin{aligned} \langle \widehat{H} \rangle &= \sum_{n=0}^{\infty} \sum_{m=0}^{\infty} c_n^* c_m E_m \delta_{nm} = \\ &= \sum_{n=0}^{\infty} c_n^* c_n E_n = \\ &= \sum_{n=0}^{\infty} |c_n|^2 E_n . \end{aligned}$$

Energia najnižšieho (základného) stavu je E_0 . Výraz trochu upravíme

$$\begin{aligned} \langle \widehat{H} \rangle &= \sum_{n=0}^{\infty} |c_n|^2 (E_n - E_0 + E_0) = \\ &= \sum_{n=0}^{\infty} |c_n|^2 (E_n - E_0) + E_0 \sum_{n=0}^{\infty} |c_n|^2 = \\ \langle \widehat{H} \rangle &= \sum_{n=0}^{\infty} |c_n|^2 (E_n - E_0) + E_0 , \end{aligned} \quad (1.48)$$

¹²V pravom slova zmysle nečasovú Schrödingerovu rovnicu ani neriešime.

¹³Pre jednoduchosť používame skrátenie a namiesto $|\psi_n(\mathbf{r}_j)\rangle$ používame $|\psi_n\rangle$.

kde sme využili to, že funkcia ψ je normovaná, a teda pre ňu platí

$$\begin{aligned}
 \langle \psi | \psi \rangle &= \left\langle \sum_{n=0}^{\infty} c_n \psi_n \left| \sum_{m=0}^{\infty} c_m \psi_m \right. \right\rangle = \\
 &= \sum_{n=0}^{\infty} \sum_{m=0}^{\infty} c_n^* c_m \langle \psi_n | \psi_m \rangle = \\
 &= \sum_{n=0}^{\infty} \sum_{m=0}^{\infty} c_n^* c_m \delta_{nm} = \\
 &= \sum_{n=0}^{\infty} c_n^* c_n = \\
 &= \sum_{n=0}^{\infty} |c_n|^2 = 1.
 \end{aligned}$$

Keďže sme v predpoklade (1.44) predpokladali usporiadanú bázu, tak v rovnici (1.48) sú všetky výrazy $|c_n|^2(E_n - E_0)$ nezáporné

$$|c_n|^2(E_n - E_0) \geq 0. \quad (1.49)$$

A tak úpravou rovnice (1.48) dostaneme

$$\begin{aligned}
 \langle \widehat{H} \rangle - E_0 &= \sum_{n=0}^{\infty} |c_n|^2(E_n - E_0) \geq 0, \\
 \langle \widehat{H} \rangle &\geq E_0.
 \end{aligned} \quad (1.50)$$

V tomto vzťahu tkvie základ hľadania vlnovej funkcie a energie základného stavu. Ústrednou myšlienkou je voľba nejakej funkcie ψ , ktorá má viacero parametrov. Variáciou parametrov meníme vlnovú funkciu aj samotnú strednú hodnotu energiu. Pomocou minimalizačných metód meníme parametre tak, aby sme sa dostali ku najnižšej hodnote, ktorá bude vždy horným odhadom energie základného stavu. Ak je funkcia volená šikovne a zároveň má dostatok parametrov pre variabilitu, tak sa dá s odhadom energie E_0 dostať bližšie. Zároveň tak vlnová funkcia ψ pre dané parametre je kandidátom na vlnovú funkciu základného stavu¹⁴.

Ďalším dôležitý ukazovateľom je stredná hodnota rozptylu od nejakého vlast-

¹⁴Funkcia ψ nebude nikdy presne vlnovou funkciou základného stavu, ale nemusí sa od nej veľmi líšiť.

ného čísla E_k . Podľa definície dostávame

$$\begin{aligned}
\langle (\widehat{H} - E_k)^2 \rangle &= \langle \psi | (\widehat{H} - E_k)^2 \psi \rangle = \\
&= \left\langle \sum_{n=0}^{\infty} c_n \psi_n \left| (\widehat{H} - E_k)^2 \sum_{m=0}^{\infty} c_m \psi_m \right. \right\rangle = \\
&= \sum_{n=0}^{\infty} \sum_{m=0}^{\infty} c_n^* c_m \langle \psi_n | (\widehat{H} - E_k)^2 \psi_m \rangle = \\
&= \sum_{n=0}^{\infty} \sum_{m=0}^{\infty} c_n^* c_m (E_m - E_k)^2 \langle \psi_n | \psi_m \rangle = \\
&= \sum_{n=0}^{\infty} \sum_{m=0}^{\infty} c_n^* c_m (E_m - E_k)^2 \delta_{nm} = \\
&= \sum_{n=0}^{\infty} |c_n|^2 (E_n - E_k)^2 \geq 0.
\end{aligned}$$

Pri úpravách sme znova použili rozklad do báze, linearitu v druhom argumente skalárneho súčinu, antilinearitu v prvom argumente skalárneho súčinu, zapôsobili sme operátorom $(\widehat{H} - E_k)^2$ na funkciu ψ_m a využili sme ortonormalitu báze. Daná hodnota je vždy väčšia alebo rovná nule.

Nás bude zaujímať prípad, kedy sa výraz rovná nule. Každý člen sčítania je nezáporný, teda každý z nich musí byť nulový. Preto ak energia E_n nejakého bázového stavu ψ_n sa nerovná energii E_k vybraného vlastného stavu ψ_k , tak musí byť koeficient $|c_n|^2$ nulový. To je splnené iba pre $c_n = 0$.

Teda výraz bude nulový, tak funkcia bude mať nulové koeficienty v bázových stavoch, ktorých energia sa nerovná E_k

$$(\forall n \in \mathbb{N}_0)((E_n \neq E_k) \Rightarrow (c_n = 0)) \quad (1.51)$$

O ostatných koeficientoch nič nevraví. Ak by sme mali nedegenerovanú energiu E_k , tak z normovacej podmienky dostávame

$$\sum_{n=0}^{\infty} |c_n|^2 = |c_k|^2 = 1. \quad (1.52)$$

Ak by sme mali l -krát degenerovanú energiu E_k , tak z normovacej podmienky dostaneme

$$\sum_{n=0}^{\infty} |c_n|^2 = \sum_{n=k}^{k+l-1} |c_n|^2 = 1. \quad (1.53)$$

1.2.6 Viriálový teorém

Ďalším dobrým overením správnosti vlastnej funkcie je viriálový teorém.

Predpokladajme, že máme nečasový Hamiltonián systému a máme nečasovú veličinu (operátor) \widehat{A} . Predpokladajme, že systém je vo vlastnom stave, a teda vlnová funkcia má tvar (1.8). Potom stredná hodnota operátora \widehat{A} je

$$\langle \widehat{A} \rangle = \langle \psi_n(\mathbf{r}_j) e^{-\frac{E_n t}{i\hbar}} | \widehat{A} \psi_n(\mathbf{r}_j) e^{\frac{E_n t}{i\hbar}} \rangle = \langle \psi_n(\mathbf{r}_j) | \widehat{A} \psi_n(\mathbf{r}_j) \rangle, \quad (1.54)$$

kde nám vypadla časová závislosť. Ale časová derivácia strednej hodnoty $\langle \hat{A} \rangle$ je

$$\frac{d}{dt} \langle \hat{A} \rangle = \frac{\partial}{\partial t} \langle \hat{A} \rangle + \frac{1}{i\hbar} \langle [\hat{A}, \hat{H}] \rangle. \quad (1.55)$$

Keďže však ani \hat{A} , ani $\langle \hat{A} \rangle$ nezávisia od času, budú ich derivácie nulové a dostávame (*hyperviriálový teorém*)

$$\langle [\hat{A}, \hat{H}] \rangle = 0. \quad (1.56)$$

Zvolíme si za operátor \hat{A} operátor $\hat{p} \cdot \hat{r}$, ktorý je pri N -časticovom systéme definovaný ako *standardný skalárny súčin*

$$\hat{p} \cdot \hat{r} = \sum_{k=1}^{3N} \hat{p}_k \hat{x}_k, \quad (1.57)$$

kde operátormi súradníc l -tej častice sú operátory \hat{x}_{3l-2} , \hat{x}_{3l-1} a \hat{x}_{3l} , operátormi hybnosti l -tej častice sú operátory \hat{p}_{3l-2} , \hat{p}_{3l-1} a \hat{p}_{3l} .

Komutátor $[\hat{p} \cdot \hat{r}, \hat{H}]$ rozpíšeme ako

$$[\hat{p} \cdot \hat{r}, \hat{H}] = \left[\sum_{k=1}^{3N} \hat{p}_k \hat{x}_k, \hat{H} \right] = \sum_{k=1}^{3N} [\hat{p}_k \hat{x}_k, \hat{H}] = \sum_{k=1}^{3N} \hat{p}_k [\hat{x}_k, \hat{H}] + [\hat{p}_k, \hat{H}] \hat{x}_k, \quad (1.58)$$

kde sme využili definíciu operátora $\hat{p} \cdot \hat{r}$, linearitu komutátora v argumentoch a vzťah

$$[\hat{A}\hat{B}, \hat{C}] = \hat{A}[\hat{B}, \hat{C}] + [\hat{A}, \hat{C}]\hat{B}, \quad (1.59)$$

ktorý môžeme overiť jednoduchým vyjadrením definície komutátora

$$\begin{aligned} [\hat{A}\hat{B}, \hat{C}] &= \hat{A}\hat{B}\hat{C} - \hat{C}\hat{A}\hat{B}, \\ \hat{A}[\hat{B}, \hat{C}] + [\hat{A}, \hat{C}]\hat{B} &= \hat{A}(\hat{B}\hat{C} - \hat{C}\hat{B}) + (\hat{A}\hat{C} - \hat{C}\hat{A})\hat{B} = \\ &= \hat{A}\hat{B}\hat{C} - \hat{A}\hat{C}\hat{B} + \hat{A}\hat{C}\hat{B} - \hat{C}\hat{A}\hat{B} = \\ &= \hat{A}\hat{B}\hat{C} - \hat{C}\hat{A}\hat{B}. \end{aligned}$$

V dodatku 4.3 sme ukázali, čomu sa komutátory $[\hat{x}_k, \hat{H}]$ a $[\hat{p}_k, \hat{H}]$ v nerelativistickom prípade rovnajú

$$\begin{aligned} [\hat{x}_k, \hat{H}] &= \frac{i\hbar}{m_k} \hat{p}_k, \\ [\hat{p}_k, \hat{H}] &= -i\hbar \frac{\partial V}{\partial x_k}. \end{aligned}$$

Dosadením do vzťahu (1.58) dostávame

$$\begin{aligned} [\hat{p} \cdot \hat{r}, \hat{H}] &= i\hbar \sum_{k=1}^{3N} \frac{\hat{p}_k \hat{p}_k}{m_k} - \frac{\partial V}{\partial x_k} \hat{x}_k = \\ &= i\hbar (2\hat{T} - \nabla V \cdot \hat{r}). \end{aligned}$$

Dosadením do vzťahu (1.56) dostávame (*virialový teorém*)

$$\langle \nabla V \cdot \hat{r} \rangle = 2 \langle T \rangle. \quad (1.60)$$

Dôležité však je, že keď máme viacčasticový systém a jediný druh interakcie medzi časticami K a L je popísaný vzájomným sféricky symetrickým potenciálom tvaru $V_{KL} = \alpha_{KL} r_{KL}^n$, tak podľa dodatku 4.4 dostaneme

$$\langle \nabla V \cdot \hat{r} \rangle = n \langle V \rangle = 2 \langle T \rangle . \quad (1.61)$$

V našom prípade používame Coulombovský potenciál, kde $n = -1$, teda dostaneme

$$- \langle V \rangle = 2 \langle T \rangle . \quad (1.62)$$

Naviac vieme strednú hodnotu energie E

$$\langle E \rangle = \langle \hat{H} \rangle = \langle \hat{T} + \hat{V} \rangle = \langle T \rangle + \langle V \rangle . \quad (1.63)$$

1.2.7 Asymptotické správanie v Coulombovskom potenciáli

Ak keď nečasovú Schrödingerovu rovnicu pre viacčasticový systém s Coulombovským potenciálom nevieme vypočítať analyticky, vieme však vyšetriť, aké asymptotické správanie musí mať vlnová funkcia v súradnicovej reprezentácii.

Predpokladajme nečasový nerelativistický Hamiltonián pre systém n častíc s Coulombovským potenciálom

$$\hat{H} := - \sum_{k=1}^n \frac{\hbar^2}{2m_k} \Delta_k + \sum_{k=1}^n \sum_{l=k+1}^n \frac{k_C Q_k Q_l}{r_{kl}} , \quad (1.64)$$

kde Q_i je náboj i -tej častice, k_C Coulombova konštanta a r_{kl} vzdialenosť k -tej a l -tej častice.

Ďalej predpokladajme, že systém je popísaný vo vlastnom súradnicovom systéme (v klasickej mechanike by to znamenalo, že sa ťažisko systému nehýbe) a teda translačný a rotačný pohyb systému sú odtransformované. Znamená to aj, že môžeme prejsť od vlnovej funkcie závislej od súradníc častíc ku vlnovej funkcii závislej od relatívnej polohy častíc. Súborom vzdialeností (r_{kl}) je v tomto prípade systém častíc úplne popísaný a môžeme teda prejsť ku novej vlnovej funkcii ψ_f závislej od jednotlivých vzdialeností častíc (r_{kl})

$$\psi(\mathbf{r}_j) = \psi_f(r_{kl}) . \quad (1.65)$$

Vlastný stav potom musí spĺňať rovnicu

$$\hat{H} \psi_{f,N}(r_{kl}) = E_N \psi_{f,N}(r_{kl}) , \quad (1.66)$$

kde na vzdialenosti r_{kl} treba pozeráť, ako na funkcie súradníc (\mathbf{r}_j).

Budeme skúmať asymptotické správanie pre $r_{kl} \rightarrow 0$ a $r_{kl} \rightarrow \infty$.

Pre väčšiu prehľadnosť je vhodné prejsť do atómových jednotiek. Hamiltonián má potom tvar

$$\hat{H} = - \sum_{k=1}^n \frac{1}{2m_k} \Delta_k + \sum_{k=1}^n \sum_{l=k+1}^n \frac{Z_k Z_l}{r_{kl}} , \quad (1.67)$$

kde m_i je hmotnosť i -tej častice v násobkoch hmotnosti elektrónu a Z_i je náboj i -tej častice v násobkoch elementárneho náboja.

Keď $r_{kl} \rightarrow 0$, tak v rovnici (1.66) je pravá strana stále konečná (lebo aj energia i vlnová funkcia sú konečné). Na ľavej strane však máme potenciál $\sim \frac{1}{r_{kl}}$, ktorý pri $r_{kl} \rightarrow 0$ ide do nekonečna. Je teda zrejmé, že niektorý iný člen/členy Hamiltoniánu túto divergenciu do nekonečna ruší/rušia. Budeme musieť vypočítať príslušné derivácie podľa súradníc (\mathbf{r}_j).

Vzdialenosť r_{kl} je definovaná ako euklidovská norma rozdielu polohových vektorov k -tej a l -tej častice

$$r_{kl} := \|\mathbf{r}_l - \mathbf{r}_k\| = \left[(x_k - x_l)^2 + (y_k - y_l)^2 + (z_k - z_l)^2 \right]^{1/2}. \quad (1.68)$$

Prvá derivácia podľa súradnice j -tej častice $a_j \in \{x, y, z\}$ potom je

$$\frac{\partial r_{kl}}{\partial a_j}(\mathbf{r}_i) = (\delta_{jk} - \delta_{jl}) \frac{a_k - a_l}{r_{kl}}. \quad (1.69)$$

Druhá derivácia podľa rovnakej súradnice je

$$\frac{\partial^2 r_{kl}}{\partial a_j^2}(\mathbf{r}_i) = (\delta_{jk} - \delta_{jl})^2 \left[\frac{1}{r_{kl}} - \frac{(a_k - a_l)^2}{r_{kl}^3} \right]. \quad (1.70)$$

Pri rátaní Laplaciánu (kinetického člena Hamiltoniánu) podľa j -tej častice potrebujeme najprv prvú deriváciu

$$\frac{\partial \psi_{f,N}}{\partial a_j}(r_{mn}) = \sum_{k=1}^n \sum_{l=k+1}^n \frac{\partial \psi_{f,N}}{\partial r_{kl}}(r_{mn}) \frac{\partial r_{kl}}{\partial a_j}(\mathbf{r}_i), \quad (1.71)$$

a potom druhú deriváciu

$$\frac{\partial^2 \psi_{f,N}}{\partial a_j^2}(r_{mn}) = \sum_{k=1}^n \sum_{l=k+1}^n \frac{\partial^2 \psi_{f,N}}{\partial r_{kl}^2}(r_{mn}) \left(\frac{\partial r_{kl}}{\partial a_j}(\mathbf{r}_i) \right)^2 + \frac{\partial \psi_{f,N}}{\partial r_{kl}}(r_{mn}) \frac{\partial^2 r_{kl}}{\partial a_j^2}(\mathbf{r}_i). \quad (1.72)$$

Dosadením vzťahov (1.69) a (1.70) do (1.72) a sčítaním cez tri súradnice x, y a z dostaneme

$$\Delta_j \psi_{f,N}(r_{mn}) = \sum_{k=1}^n \sum_{l=k+1}^n (\delta_{jk} - \delta_{jl})^2 \left[\frac{\partial^2 \psi_{f,N}}{\partial r_{kl}^2}(r_{mn}) + \frac{\partial \psi_{f,N}}{\partial r_{kl}}(r_{mn}) \frac{2}{r_{kl}} \right]. \quad (1.73)$$

Presčítaním, preindexovaním a dosadením do (1.66) dostaneme

$$\begin{aligned} - \sum_{j=1}^n \frac{1}{2m_j} \sum_{\substack{k=1 \\ k \neq j}}^n \left[\frac{\partial^2 \psi_{f,N}}{\partial r_{jk}^2}(r_{mn}) + \frac{\partial \psi_{f,N}}{\partial r_{jk}}(r_{mn}) \frac{2}{r_{jk}} \right] + \sum_{j=1}^n \sum_{k=j+1}^n \frac{Z_j Z_k}{r_{jk}} \psi_{f,N}(r_{mn}) = \\ = E_N \psi_{f,N}(r_{mn}). \end{aligned}$$

Ďalej budeme predpokladať, že funkcia $\psi_{f,N}$ má všade konečné prvé i druhé derivácie. Keď budeme pokračovať v pôvodnej analýze a vyberieme si konkrétnu vzdialenosť r_{jk} , tak aby sme zabránili divergovaniu ľavej strane rovnice, budú musieť výrazy pri $\frac{1}{r_{jk}}$ ísť limitne k nule s minimálnym poriadkom r_{jk} . Tým dostaneme rovnicu

$$\lim_{r_{jk} \rightarrow 0} \left[- \left(\frac{1}{m_j} + \frac{1}{m_k} \right) \frac{\partial \psi_{f,N}}{\partial r_{jk}}(r_{mn}) + Z_j Z_k \psi_{f,N}(r_{mn}) \right] = 0. \quad (1.74)$$

Zavedením redukovanej hmotnosti $\frac{1}{\mu_{jk}} = \frac{1}{m_j} + \frac{1}{m_k}$ a upravením limity dostaneme

$$\lim_{r_{jk} \rightarrow 0} \frac{\frac{\partial \psi_{f,N}}{\partial r_{jk}}}{\psi_{f,N}}(r_{mn}) = Z_j Z_k \mu_{jk}. \quad (1.75)$$

Ak by sme predpokladali rovnosť pomeru v limite v okolí, tak integrovaním dostaneme závislosť vlnovej funkcie

$$\psi_{f,N}(r_{mn}) \sim e^{Z_j Z_k \mu_{jk} r_{jk}}. \quad (1.76)$$

Podobne sa dá uvažovať aj o asymptotickom správaní pri limite $r_i \rightarrow \infty$. Vtedy je jedna i -tá častica poslaná do nekonečnej vzdialenosti. Budeme vychádzať z rovnice (1.66) s Hamiltoniánom z (1.67). Najprv rozdelíme Hamiltonián na členy častíc, ktoré v systéme ostávajú do \widehat{H}_{zV} a členy Hamiltoniánu i -tej častice

$$-\frac{1}{2m_i} \Delta_i \psi_{f,N}(r_{mn}) + \sum_{\substack{k=1 \\ k \neq i}}^n \frac{Z_i Z_k}{r_{ik}} \psi_{f,N}(r_{mn}) + \widehat{H}_{zV} \psi_{f,N}(r_{mn}) = E_N \psi_{f,N}(r_{mn}). \quad (1.77)$$

Bude nás zaujímať limitné správanie tejto rovnice pre $r_i \rightarrow \infty$. Vieme, že ak je i -tá častica ďaleko, tak sa ide o izolovaný systém zvyšných častíc, ktoré majú svoje stavy a energie. Teda pre $r_i \rightarrow \infty$ sa $\widehat{H}_{zV} \psi_{f,N}(r_{mn})$ správa ako $\tilde{E}_N \psi_{f,N}(r_{mn})$, kde \tilde{E}_N má význam energie systému zvyšných častíc. Keďže máme viazaný stav, tak $\psi_{f,N}(r_{mn})$ ide s určitým trendom pre $r_i \rightarrow \infty$ do nuly. Lenže zatiaľ čo členy $\widehat{H}_{zV} \psi_{f,N}(r_{mn})$ a $E_N \psi_{f,N}(r_{mn})$ klesajú rovnako rýchlo ako $\psi_{f,N}(r_{mn})$, len sú pre násobené konštantou, tak člen $\frac{Z_i Z_k}{r_{ik}} \psi_{f,N}(r_{mn})$ klesá o jeden rád rýchlejšie. Preto ho môžeme v limitnom tvare rovnice vynechať. Dostaneme teda

$$-\frac{1}{2m_i} \Delta_i \psi_{f,N}(r_{mn}) + (\tilde{E}_N - E_N) \psi_{f,N}(r_{mn}) = 0. \quad (1.78)$$

Uvedená rovnica je pohybová rovnica voľnej častice. Keď vyberieme riešenie, ktoré klesá k nule, dostaneme približnú závislosť

$$\psi_{f,N}(r_{mn}) \sim e^{-\sqrt{2m_i I} r_i}, \quad (1.79)$$

kde I sme označili rozdiel energií \tilde{E}_N a E_N s významom ionizačnej energie.

1.3 Bornova-Oppenheimerova aproximácia

Predpokladajme nečasový nerelativistický Hamiltonián pre systém N jadier a n elektrónov

$$\widehat{H} := - \sum_{k=1}^n \frac{\hbar^2}{2m_k} \Delta_k - \sum_{l=1}^N \frac{\hbar^2}{2M_l} \Delta_l + V(\mathbf{r}_j, \mathbf{R}_j), \quad (1.80)$$

kde malými veličinami označujeme veličiny elektrónov (m_k, \mathbf{r}_j) a veľké veličiny sú veličiny jadier (M_l, \mathbf{R}_j).

Nečasová Schrödingerova rovnica potom má tvar

$$\widehat{H} \psi(\mathbf{r}_j, \mathbf{R}_j) = E \psi(\mathbf{r}_j, \mathbf{R}_j). \quad (1.81)$$

Keďže hmotnostný rozdiel medzi jadrami a elektrónmi je niekoľko rádov (3 až 5), tak môžeme uvažovať, že elektróny sa pohybujú omnoho rýchlejšie. Teda

elektróny sa pohybujú v poli nehybných jadier a vzájomnej interakcie elektrónov. Zato jadrá reagujú na vystredované pole vytvárané elektrónmi a pole od ostatných jadier. To však zodpovedá dvom Schrödingerovým rovniciam, osobitne pre elektróny a osobitne jadrá. Tým pádom je ich vzájomný pohyb nezávislý a to znamená, že vlnovú funkciu systému dostaneme súčinom vlnových funkcií elektrónov ψ_e a jadier ψ_J

$$\psi(\mathbf{r}_j, \mathbf{R}_j) = \psi_e(\mathbf{r}_j; \mathbf{R}_j)\psi_J(\mathbf{R}_j), \quad (1.82)$$

kde si musíme všimnúť, že vlnová funkcia jadier neobsahuje súradnice elektrónov, lebo sa pohybuje v ich strednom poli. Naopak vo vlnovej funkcii elektrónov súradnice jadier vystupujú, lebo elektróny sa pohybujú v poli nehybných jadier v daných polohách, ale treba upozorniť, že už tu vo vlnovej funkcii majú charakter parametrov a nie premenných.

Nečasová Schrödingerova rovnica pre elektróny potom podľa vyslovených predpokladov má tvar

$$\left(-\sum_{k=1}^n \frac{\hbar^2}{2m_k} \Delta_k + V(\mathbf{r}_j, \mathbf{R}_j) \right) \psi_e(\mathbf{r}_j; \mathbf{R}_j) = U(\mathbf{R}_j)\psi_e(\mathbf{r}_j; \mathbf{R}_j). \quad (1.83)$$

Keď vezmeme predpoklad o vlnovej funkcii (1.82) a dosadíme ho do (1.81), dostaneme

$$\left(-\sum_{k=1}^n \frac{\hbar^2}{2m_k} \Delta_k - \sum_{l=1}^N \frac{\hbar^2}{2M_l} \Delta_l + V(\mathbf{r}_j, \mathbf{R}_j) \right) \psi_e(\mathbf{r}_j; \mathbf{R}_j)\psi_J(\mathbf{R}_j) = E\psi_e(\mathbf{r}_j; \mathbf{R}_j)\psi_J(\mathbf{R}_j). \quad (1.84)$$

Využitím rovnice (1.83) dostaneme

$$U(\mathbf{R}_j)\psi_e(\mathbf{r}_j; \mathbf{R}_j)\psi_J(\mathbf{R}_j) - \sum_{l=1}^N \frac{\hbar^2}{2M_l} [\Delta_l \psi_e(\mathbf{r}_j; \mathbf{R}_j)\psi_J(\mathbf{R}_j) + 2\nabla_l \psi_e(\mathbf{r}_j; \mathbf{R}_j)\nabla_l \psi_J(\mathbf{R}_j) + \psi_e(\mathbf{r}_j; \mathbf{R}_j)\Delta_l \psi_J(\mathbf{R}_j)] = E\psi_e(\mathbf{r}_j; \mathbf{R}_j)\psi_J(\mathbf{R}_j).$$

Táto rovnica predstavuje *adiabatickú aproximáciu* riešenia.

Keď však ešte k tomu pridáme predpoklady, že vlnová funkcia elektrónov sa so zmenou polohy jadier veľmi nemení (derivácie funkcie ψ_e podľa \mathbf{R}_j sú malé)

$$\Delta_l \psi_e(\mathbf{r}_j; \mathbf{R}_j) = 0 \quad \nabla_l \psi_e(\mathbf{r}_j; \mathbf{R}_j) = 0,$$

tak dostaneme zjednodušenú rovnicu

$$\left(-\sum_{l=1}^N \frac{\hbar^2}{2M_l} \Delta_l + U(\mathbf{R}_j) \right) \psi_J(\mathbf{R}_j) = E\psi_J(\mathbf{R}_j), \quad (1.85)$$

ktorá sa nazýva *Bornova-Oppenheimerova aproximácia*. Je badateľné, že rovnica má tvar, ako keby išlo o nečasovú Schrödingerovú rovnicu jadier v potenciáli vytváranom interakciou jadier a energiou elektrónov.

2. Výpočtové algoritmy

2.1 Monte Carlo metóda

Základná myšlienka využitia variačného princípu bola vyslovená už v odseku 1.2.5. Jediným problémom je, že hodnota $\langle \widehat{H} \rangle$, ktorú je pre danú vlnovú funkciu ψ treba vypočítať, sa nedá priamočiaro dostať vždy. $\langle \widehat{H} \rangle$ je z definície integrál a sa dá jednoducho analyticky vyrátať len pre určité vlnové funkcie. Tu prichádzajú na pomoc algoritmy skupiny *Monte Carlo*. Tieto algoritmy sa používajú na rátanie integrálov pomocou generovania náhodných čísel. Problém rátania integrálu sa prevedie náhodné generovanie veličiny, ktorej stredná hodnota zodpovedá hľadanej hodnote integrálu.

Predstavme si, že máme rátať daný integrál

$$\langle A \rangle = \int_V \varrho(\mathbf{r}) A(\mathbf{r}) dV, \quad (2.1)$$

kde máme dve funkcie ϱ a A . Funkcia $\varrho(\mathbf{r})$ má význam hustoty pravdepodobnosti na množine V . Preto pre túto funkciu musí platiť podmienka

$$\int_V \varrho(\mathbf{r}) dV = 1, \quad (2.2)$$

ktorá vyjadruje, že pravdepodobnosť výberu z celej množiny je jednotková (istá). Funkcia $A(\mathbf{r})$ tak nadobúda význam veličiny, ktorej stredná hodnota sa na rozdelení $\varrho(\mathbf{r})$ ráta. Potom je už z tvaru integrálu jasné, že stredná hodnota veličiny $A(\mathbf{r})$ na rozdelení $\varrho(\mathbf{r})$ je práve hľadaná hodnota $\langle A \rangle$.

Preto ak budeme generovať vzorky \mathbf{r}_j z rozdelenia $\varrho(\mathbf{r})$ a budeme pre ne rátať hodnoty veličiny $A(\mathbf{r})$

$$A_j := A(\mathbf{r}_j), \quad (2.3)$$

tak tieto vzorky sú z nejakého rozdelenia, ktoré má svoju strednú hodnotu μ_A a rozptyl σ_A^2

$$\mathcal{E}(A) = \mu_A, \quad (2.4)$$

$$\mathcal{V}ar(A) = \sigma_A^2. \quad (2.5)$$

Stredná hodnota μ_A sa rovná hodnote $\langle A \rangle$

$$\mu_A = \langle A \rangle. \quad (2.6)$$

Dobrym odhadom strednej hodnoty veličiny A zo vzoriek A_j je *výberový priemer*

$$\tilde{A} := \frac{1}{N} \sum_{j=1}^N A_j, \quad (2.7)$$

kde N je počet vzoriek výberu. Vlastnosťami výberového priemeru sú

$$\mathcal{E}(\tilde{A}) = \mu_A, \quad (2.8)$$

$$\mathcal{V}ar(\tilde{A}) = \frac{\sigma_A^2}{N}. \quad (2.9)$$

Samozrejme, okrem odhadu hľadanej hodnoty potrebujeme vedieť aj odhad chyby odhadu. Ako chyba sa dobre volí $\sqrt{\text{Var}(\tilde{A})}$, lenže hodnota σ_A^2 daného rozdelenia je nám neznáma, a preto ju budeme musieť tiež získať z výberu vzoriek A_j . Dobrým odhadom rozptylu σ_A^2 je

$$\tilde{\sigma}_A^2 := \frac{1}{N} \sum_{j=1}^N (A_j - \mu_A)^2, \quad (2.10)$$

pretože jeho vlastnosťou je

$$\mathcal{E}(\tilde{\sigma}_A^2) = \sigma_A^2. \quad (2.11)$$

Ale problémom je, že aj hodnota μ_A daného rozdelenia je nám neznáma (veď práve túto hodnotu hľadáme), preto ako odhad rozptylu σ_A^2 použijeme estimátor

$$\check{\sigma}_A^2 := \frac{1}{N} \sum_{j=1}^N (A_j - \tilde{A})^2, \quad (2.12)$$

lenže jeho vlastnosťou je

$$\mathcal{E}(\check{\sigma}_A^2) = \frac{N-1}{N} \sigma_A^2. \quad (2.13)$$

Preto finálnym odhadom rozptylu σ_A^2 je estimátor

$$\hat{\sigma}_A^2 := \frac{1}{N-1} \sum_{j=1}^N (A_j - \tilde{A})^2, \quad (2.14)$$

ktorý už nie je vychýlený a jeho vlastnosťou je

$$\mathcal{E}(\hat{\sigma}_A^2) = \sigma_A^2. \quad (2.15)$$

Potom odhadom smerodajnej odchýlky výberového priemeru (teda odhad chyby vypočítanej hodnoty) je

$$\Delta \tilde{A} := \sqrt{\frac{1}{N(N-1)} \sum_{j=1}^N (A_j - \tilde{A})^2}. \quad (2.16)$$

Vzťahy (2.8), (2.9), (2.11), (2.13) a (2.15) sme vypočítali v dodatku 4.5 za predpokladu nezávislosti vzoriek A_j .

2.1.1 Metropolisov-Hastingsov algoritmus

Účelom Metropolisovho-Hastingsovho algoritmu je zabezpečiť výber vzoriek \mathbf{r}_j z daného rozdelenia $\varrho(\mathbf{r})$. Algoritmus využíva *Markovov proces*¹, ktorý konverguje ku výberu z daného rozdelenia $\varrho(\mathbf{r})$. Markovov proces je daný pravdepodobnosťami $P(\mathbf{r} \rightarrow \mathbf{r}')$ prechodu od stavu (vzorky) \mathbf{r} do stavu (vzorky) \mathbf{r}' . Aby platilo, že daný Markovov proces sa asymptoticky blíži rozdeleniu $\varrho(\mathbf{r})$, musí platiť

$$\varrho(\mathbf{r})P(\mathbf{r} \rightarrow \mathbf{r}') = \varrho(\mathbf{r}')P(\mathbf{r}' \rightarrow \mathbf{r}). \quad (2.17)$$

¹Rozhodovací proces výberu nového stavu v závislosti na predchádzajúcom stave a vybranou akciou.

Pravdepodobnosť prechodu od stavu \mathbf{r} do stavu \mathbf{r}' môžeme rozdeliť na pravdepodobnosť výberu $V(\mathbf{r} \rightarrow \mathbf{r}')$ stavu \mathbf{r}' zo stavu \mathbf{r} a pravdepodobnosť prijatia novej pozície $A(\mathbf{r} \rightarrow \mathbf{r}')$

$$P(\mathbf{r} \rightarrow \mathbf{r}') = V(\mathbf{r} \rightarrow \mathbf{r}')A(\mathbf{r} \rightarrow \mathbf{r}'). \quad (2.18)$$

Potom dosadením dostávame podmienku pre pravdepodobnosť prijatia

$$\varrho(\mathbf{r})V(\mathbf{r} \rightarrow \mathbf{r}')A(\mathbf{r} \rightarrow \mathbf{r}') = \varrho(\mathbf{r}')V(\mathbf{r}' \rightarrow \mathbf{r})A(\mathbf{r}' \rightarrow \mathbf{r}), \quad (2.19)$$

$$\frac{A(\mathbf{r} \rightarrow \mathbf{r}')}{A(\mathbf{r}' \rightarrow \mathbf{r})} = \frac{\varrho(\mathbf{r}')V(\mathbf{r}' \rightarrow \mathbf{r})}{\varrho(\mathbf{r})V(\mathbf{r} \rightarrow \mathbf{r}')}. \quad (2.20)$$

Hustota pravdepodobnosti $\varrho(\mathbf{r})$ je daná, spôsob, ako vyberáme nové vzorky zo starých $V(\mathbf{r} \rightarrow \mathbf{r}')$ si môžeme voliť. Výber tohto nám určuje, ako musí vyzerat $A(\mathbf{r} \rightarrow \mathbf{r}')$. Najčastejšia voľba, aby spĺňala predchádzajúci vzťah, je nasledujúca

$$A(\mathbf{r} \rightarrow \mathbf{r}') = \min \left\{ 1, \frac{\varrho(\mathbf{r}')V(\mathbf{r}' \rightarrow \mathbf{r})}{\varrho(\mathbf{r})V(\mathbf{r} \rightarrow \mathbf{r}')} \right\}. \quad (2.21)$$

Celý postup algoritmu teda vyzerá takto:

1. Vyberieme si náhodnú iniciačnú polohu \mathbf{r}_0 .
2. Vyberieme si možnú novú polohu \mathbf{r}'_i , ktorej pravdepodobnosť výberu je $V(\mathbf{r}_i \rightarrow \mathbf{r}'_i)$.
3. Vypočítame hodnotu $p = \frac{\varrho(\mathbf{r}'_i)V(\mathbf{r}'_i \rightarrow \mathbf{r}_i)}{\varrho(\mathbf{r}_i)V(\mathbf{r}_i \rightarrow \mathbf{r}'_i)}$ a vyberieme náhodné číslo a zo spojitého uniformného rozdelenia $U(0, 1)$.
4. Ak $a < p$, tak nová poloha je $\mathbf{r}_{i+1} = \mathbf{r}'_i$, inak nová poloha je $\mathbf{r}_{i+1} = \mathbf{r}_i$.
5. Vrátime sa k bodu 2 a proces opakujeme.

2.2 Závislosť vyberaných vzoriek

Zatiaľ čo v dodatku 4.5 sme pri rátaní estimátorov predpokladali nezávislosť vzoriek A_j , tak z Metropolisovho-Hastingsovho algoritmu generujeme vzorky, ktoré budú závislé. Síce s väčším odstupom sa vzorky postupne stávajú nezávislé, tak táto závislosť nám spôsobuje nesprávny odhad chyby. Chyba výsledkov je potom výrazne podhodnotená a preto algoritmy závislé na tejto chybe môžu vyhodnotiť niektoré kroky nesprávne. Preto sme pre výpočet chyby použili *metódu blokovania*.

2.2.1 Metóda blokovania

Metóda blokovania [5] je metóda na rýchly odhad chyby korelovaných vzoriek. Z Metropolisovho-Hastingsovho algoritmu máme n vzoriek A_j . V jednom kroku metódy blokovania transformujeme dáta A_j do súboru dát A'_j polovičnej veľkosti $n' = \frac{n}{2}$. Transformácia je daná predpisom

$$A'_j = \frac{A_{2j-1} + A_{2j}}{2}, \quad (2.22)$$

²Predpokladajme, že máme 2^k vzoriek, vtedy nie je problém so zahadzovaním posledného prvku nepárneho počtu dát.

Ukazuje sa, že pri tejto transformácii sa zachováva stredná hodnota výberového priemeru $\mathcal{E}(\tilde{A})$ i rozptyl výberového priemeru $\mathcal{Var}(\tilde{A})$. V prípade nezávislých vzoriek je odhad $\check{\sigma}_A^2$ správny. V prípade závislých vzoriek je podhodnotený, preto stále platí

$$\mathcal{Var}(\tilde{A}) \geq \check{\sigma}_A^2. \quad (2.23)$$

Po jednom kroku transformácie sa odhad $\check{\sigma}_A^2$ pri závislých vzorkách zvýši, pri nezávislých nezmení. Proces transformácie sa opakuje, až kým nie je dosiahnutá konštantná hodnota odhadu $\check{\sigma}_A^2$ na "novej" množine vzoriek. Podľa [5] aj s odhadom chyby má odhad $\mathcal{Var}(\tilde{A})$ tvar

$$\mathcal{Var}(\tilde{A}) \approx \check{\sigma}_A^2 \left(1 \pm \sqrt{\frac{2}{n-1}} \right), \quad (2.24)$$

keď je dosiahnutá konštantná hodnota.

Práve odhad chyby je dobrým ukazovateľom, ktorý pomôže určiť, kedy máme transformáciu blokovania ukončiť a tým dosiahnuť lepší odhad chyby výberového priemeru korelovaných vzoriek.

2.3 Optimalizačné metódy

Ako sme už spomínali vyššie pri variačnom princípe, základnou úlohou je minimalizácia energie stavu opísaného vlnovou funkciou ψ voči jej parametrom. Výpočet energie stavu pre určité parametre má na starosti Monte Carlo metóda. Samotný problém hľadania extrému (v našom prípade minima) funkcie viacerých premenných spadá pod rozvinutú oblasť výpočtovej techniky, ktorá sa zaoberá *optimalizáciou*. Optimalizácia v matematike znamená výber najlepšieho prvku z danej množiny vzhľadom na určité kritéria. V užšom zmysle (v jednoduchom prípade) znamená hľadanie extrému (maxima alebo minima) reálnej funkcie na nejakej množine.

Zatiaľ čo matematika nám podáva podmienky existencie extrému, nutné a postačujúce podmienky pre extrém, pre programovanie boli vyvinuté široké triedy algoritmov na výpočtové riešenie problému.

Ku každému zadaniu treba pristupovať individuálne. Niektoré metódy sú efektívnejšie, niektoré sa ani nedajú použiť.

Najprv sa budeme venovať tejto problematike v prípade optimalizácie na jednorozmernom priestore.

2.3.1 Optimalizácia v jednorozmernom priestore

V najjednoduchšom prípade hľadáme extrém reálnej funkcie na jednorozmernom priestore (optimalizácia podľa jedného parametra). Základná úloha je formulovaná nasledovne.

Máme daný otvorený interval I^3 , na ktorom je definovaná reálna konečná funkcia $F(x)$

$$F : I \rightarrow \mathbb{R}.$$

³V tomto prípade nešpecifikujeme, či je konečný, či nekonečný.

Nájdite extrém tejto funkcie na intervale I .

Samozrejme, bližšia špecifikácia vlastností funkcie $F(x)$ nám umožňuje vybrať efektívnu metódu, keďže to nám umožňuje zúžiť riešenie všeobecne zadanej úlohy na menšiu triedu funkcií rýchlejším spôsobom. V našom prípade si vystačíme s dvoma metódami: *Newtonova metóda* a *metóda zlatého rezu*.

2.3.1.1 Newtonova metóda

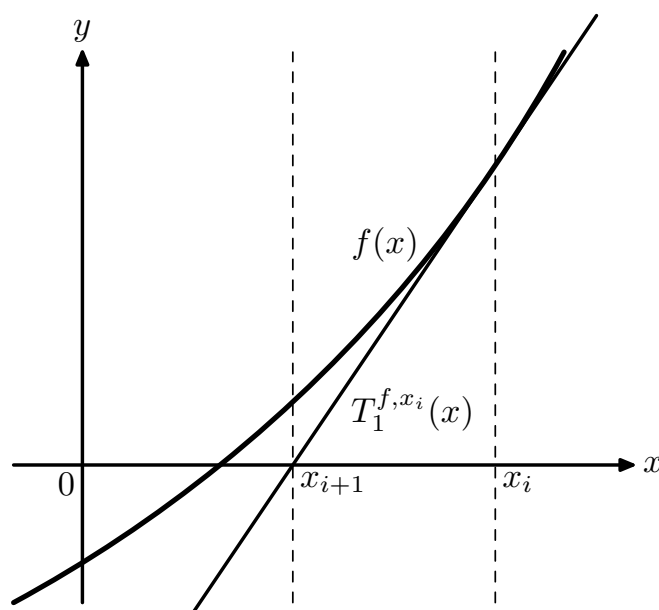
Newtonova metóda je metóda na hľadanie koreňov nejakej reálnej funkcie $f(x)$ na danom intervale I (hľadanie $x \in I$, že platí $f(x) = 0$). Je založená na aproximácii funkcie $f(x)$ Taylorovým radom 1. stupňa $T_1^{f,x_i}(x)$, ktorého koreň sa volí ako nový bod, v ktorom sa znova určuje aproximácia. Stupeň Taylorovho radu je najmenší, pri ktorom má aproximovaná funkcia koreň. Taylorov rad 1. stupňa funkcie $f(x)$ v bode x_i má tvar

$$T_1^{f,x_i}(x) = f(x_i) + f^{(1)}(x_i)(x - x_i), \quad (2.25)$$

kde $f^{(1)}(x_i)$ je prvá derivácia funkcie $f(x)$ v bode x_i . Riešenie tejto rovnice sa stáva novým bodom

$$x_{i+1} = x_i - \frac{f(x_i)}{f^{(1)}(x_i)}. \quad (2.26)$$

Na obrázku 2.1 vidíme grafické zobrazenie jedného kroku Newtonovej metódy pri hľadaní koreňa funkcie $f(x)$.



Obr. 2.1: Jeden krok Newtonovej metódy

Ako však táto metóda súvisí s hľadaním extrému? V prípade, že funkcia $F(x)$ je triedy $\mathcal{C}^2(I)$, tak nutnou podmienkou pre extrém je nulová derivácia

$$\frac{dF}{dx}(x_m) = 0, \quad (2.27)$$

kde x_m je bod, kde je extrém funkcie. A práve ak sa stotožnia funkcie $\frac{dF}{dx}(x)$ a $f(x)$, tak sa dá Newtonova metóda použiť na hľadanie extrému funkcie. Preformulujeme Newtonovu metódu pri hľadaní extrému.

Pri optimalizácii je Newtonova metóda založená na aproximácii funkcie $F(x)$ Taylorovým radom 2. stupňa $T_2^{F,a}(x)$, ktorého extrém sa volí ako nový bod, v ktorom sa znova určuje aproximácia. Stupeň Taylorovho radu je najmenší, pri ktorom má aproximovaná funkcia extrém. Taylorov rad 2. stupňa funkcie $F(x)$ v bode x_i má tvar

$$T_2^{F,x_i}(x) = F(x_i) + F^{(1)}(x_i)(x - x_i) + \frac{F^{(2)}(x_i)}{2}(x - x_i)^2. \quad (2.28)$$

Pre extrém tejto kvadratickej rovnice platí

$$\frac{dT_2^{F,x_i}}{dx}(x_{i+1}) = 0.$$

Riešenie má potom tvar

$$x_{i+1} = x_i - \frac{F^{(1)}(x_i)}{F^{(2)}(x_i)}. \quad (2.29)$$

Celý postup algoritmu potom vyzerá nasledovne:

1. Vyberieme si počiatočnú polohu x_0 .
2. Vypočítame prvú a druhú deriváciu funkcie $F(x)$ v bode x_i .
3. Podľa (2.29) vypočítame novú polohu x_{i+1} .
4. Vrátime sa k bodu 2. a proces opakujeme.

2.3.1.2 Metóda zlatého rezu

Metóda zlatého rezu je algoritmus zameraný na hľadanie extrému funkcie jednej premennej za určitých predpokladov, ktorý sa využíva pri určitých prípadoch.

Budeme sa zaoberať problémom hľadania minima, čo je náš problém.⁴ Predpokladáme, že máme daný interval $\langle a, b \rangle$, na ktorom máme známu funkciu $F(x)$

$$F : \langle a, b \rangle \rightarrow \mathbb{R},$$

ktorá je navyše na $\langle a, b \rangle$ spojitá a na intervale $\langle a, b \rangle$ má práve jedno lokálne minimum

$$(\exists! c \in \langle a, b \rangle)(\exists \delta \in \mathbb{R})(\delta > 0)(\forall x \in P_\delta(c))(F(x) \geq F(c)),$$

kde $P_\delta(c)$ je δ -prstencové okolie bodu c na intervale $\langle a, b \rangle$

$$P_\delta(c) := \{x \in \langle a, b \rangle \mid 0 < |x - c| < \delta\}.$$

Ak by funkcia bola daná analyticky a diferencovateľná na celom intervale, tak by sme minimum hľadali cez *nutnú podmienku existencie lokálneho extrému* (nulová derivácia v bode extrému). Tento algoritmus sa však používa v prípade, keď nepoznáme deriváciu funkcie (nevieme, či je vôbec funkcia diferencovateľná) a je

⁴V prípade hľadania maxima budú kroky analogické, len by boli zmenené nerovnosti.

časovo náročné vypočítať funkčnú hodnotu v jednom danom bode s dostatočnou presnosťou.

Pri hľadaní minima na základe znalosti funkčných hodnôt spojitej funkcie iba v bodoch sa v každom kroku znižuje interval, na ktorom sa minimum nachádza. Opakovaním procesu sa postupne lokalizuje s určitou chybou minimum funkcie. Na dosiahnutie danej chyby určenia minima (či už chyby funkčnej hodnoty v minime, či chyby premennej) stačí konečný počet opakovaní.

Na základe znalosti troch funkčných hodnôt pre tri premenné sa nedá interval zmenšiť. Preto je nutné poznať aspoň 4 funkčné hodnoty.

V jednom kroku teda na základe 4 funkčných hodnôt zmenšíme interval na menší podinterval, kde sa nachádza minimum. Poradové číslo kroku označíme indexom i . Interval, kde sa nachádza minimum pred i -tým krokom, označíme $\langle a_i, b_i \rangle$. Vybratý podinterval potom bude $\langle a_{i+1}, b_{i+1} \rangle$. Tieto 4 body, v ktorých určíme funkčné hodnoty označíme $x_{i,1}$, $x_{i,2}$, $x_{i,3}$ a $x_{i,4}$ a predpokladáme, že sú usporiadané a navzájom rôzne a prirodzene okrajové hodnoty sú rovné bodom na koncoch intervalu $\langle a_i, b_i \rangle$

$$a_i = x_{i,1} < x_{i,2} < x_{i,3} < x_{i,4} = b_i .$$

Ako interval $\langle a_{i+1}, b_{i+1} \rangle$ teda môžu byť vybrané dva intervaly: $\langle x_{i,1}, x_{i,3} \rangle$ alebo $\langle x_{i,2}, x_{i,4} \rangle$. Toto rozhodnutie je založené na funkčných hodnotách v týchto bodoch a vychádza z predpokladov. Zatiaľ sme popisali všeobecnú metódu pri hľadaní minima delením intervalu na 3 úseky. Vo väčšej všeobecnosti táto metóda by mohla deliť interval na viac úsekov, čím by teoreticky mohla rýchlejšie zmenšiť interval (lebo by možné intervaly boli menšie), ale bolo by to na úkor výpočtového času (viac delení, výpočet väčšieho množstva funkčných hodnôt). Tu si stačí uvedomiť, že nie je potrebné rátať nadbytočné informácie, lebo tie nás stoja výpočtový čas. Ak by sme vypočítali 5 hodnôt, v skutočnosti by bola šanca (približne 2 z 3 prípadov), že jedna z vypočítaných hodnôt sa nevyužije, čím sa stratili výpočtový čas. Tu si môžeme pomôcť analógiou v prípade inej metódy, ktorá nie je až tak vzdialená našej a síce *metódy bisekcie*. V prípade tejto metódy sa rieši podobná úloha. Máme daný interval $\langle a, b \rangle$, na ktorom máme známu funkciu $f(x)$, ktorá je navyše na $\langle a, b \rangle$ spojitá a na intervale $\langle a, b \rangle$ má práve jeden koreň (t.j. $f(x) = 0$). Vidíme, že ak by sme poznali deriváciu funkcie $F(x)$ na intervale $\langle a, b \rangle$ a tá by bola spojitá na intervale $\langle a, b \rangle$, tak by sme náš problém riešili metódou bisekcie práve pre deriváciu funkcie $F(x)$. Rovnako aj v tomto prípade by sme mohli rozdeliť interval na viac úsekov, ale rátať by sme tak redundantné informácie o funkcii $f(x)$, keďže by sme aj tak v ďalšom kroku použili iba hodnoty z krajov vybraného intervalu. Ak teda uvažujeme analogicky, je zbytočné rátať viac hodnôt ako treba na rozhodnutie na výber podintervalu. Samozrejme túto domnienku založenú iba na analógii treba podložiť konkrétnymi výpočtami, v tomto texte sa však budeme zaoberať už len skonštatujeme závery výberu optimálneho delenia.

Práve optimálny spôsob výberu počtu delení intervalu a výberu hodnôt $x_{i,2}$ a $x_{i,3}$ dal názov tejto metóde. Ukazuje sa, že ako sme predpokladali na základe analógie, že optimálne je deliť interval na 3 úseky (poznať 4 funkčné hodnoty) a interval je výhodné deliť v pomere zlatého rezu. Toto delenie má viacero vlastností, ktoré by sa dali predpokladať. V prvom rade je toto delenie symetrické voči stredu intervalu (i keď niekedy práve porušenie symetrie môže spôsobiť lepšiu voľbu). Ďalej algoritmus šetrí výpočtový čas, lebo v jednom kroku netreba rátať

2 funkčné hodnoty v strede intervalu, ale iba jednu. Práve voľba pomeru zlatého rezu zabezpečuje, že jeden z novo vybraných bodoch $x_{i,2}$ a $x_{i,3}$ padne práve na už vypočítanú hodnotu. Ďalšou vlastnosťou je, že delenie je pomerné. Body v ktorých sa rátaajú funkčné hodnoty sú na intervale $\langle a_i, b_i \rangle$ rozdelené podľa pomeru vzdialeností.

Predpokladajme, že v jednom kroku vyberieme podinterval $\langle x_{i,1}, x_{i,3} \rangle$ a chceme využiť hodnotu v bode $x_{i,2}$. Preto bod $x_{i,2}$ musí byť zároveň bodom $x_{i+1,3}$. Keďže sme vraveli, že delenie je pomerné, tak musí platiť, že pomer dĺžok intervalu $\langle x_{i,1}, x_{i,3} \rangle$ a $\langle x_{i,1}, x_{i,4} \rangle$ sa rovná pomeru intervalu $\langle x_{i+1,1}, x_{i+1,3} \rangle$ a $\langle x_{i+1,1}, x_{i+1,4} \rangle$

$$\frac{x_{i,3} - x_{i,1}}{x_{i,4} - x_{i,1}} = \frac{x_{i+1,3} - x_{i+1,1}}{x_{i+1,4} - x_{i+1,1}} = \frac{x_{i,2} - x_{i,1}}{x_{i,3} - x_{i,1}}. \quad (2.30)$$

Podobne budeme uvažovať, ak by sme vybrali podinterval $\langle x_{i,2}, x_{i,4} \rangle$. Potom dostaneme pomer

$$\frac{x_{i,4} - x_{i,2}}{x_{i,4} - x_{i,1}} = \frac{x_{i+1,4} - x_{i+1,2}}{x_{i+1,4} - x_{i+1,1}} = \frac{x_{i,4} - x_{i,3}}{x_{i,4} - x_{i,2}}. \quad (2.31)$$

Keďže sme vraveli, že delenie je pomerné, tak prejdeme od neznámych $x_{i,1}$ až $x_{i,4}$ ku neznámym X_1 až X_4 daných vzťahom

$$X_j := \frac{x_{i,j} - x_{i,1}}{x_{i,4} - x_{i,1}}. \quad (2.32)$$

Podľa definície je jasné, že $X_1 = 0$ a $X_4 = 1$. Rovnice (2.30) a (2.31) potom v nových neznámych majú tvar

$$\begin{aligned} X_3 &= \frac{X_2}{X_3}, \\ 1 - X_2 &= \frac{1 - X_3}{1 - X_2}. \end{aligned}$$

Tieto rovnice majú 4 riešenia a síce

$$\begin{aligned} X_2 &= 0, X_3 = 0; \\ X_2 &= 1, X_3 = 1; \\ X_2 &= -1 - \varphi, X_3 = -1 - \varphi; \\ X_2 &= 1 - \varphi, X_3 = \varphi; \end{aligned}$$

kde sme zaviedli označenie $\varphi = \frac{\sqrt{5}-1}{2} \approx 0,618034\dots$. Je jasné, že prvé tri riešenia sú v rozpore s rôznorodosťou hodnôt $x_{i,1}$ až $x_{i,4}$. Tretie riešenie dokonca vyberá body mimo intervalu. Jediné rozumné riešenie zodpovedá deleniu intervalu podľa zlatého rezu, podľa ktorého je metóda pomenovaná.

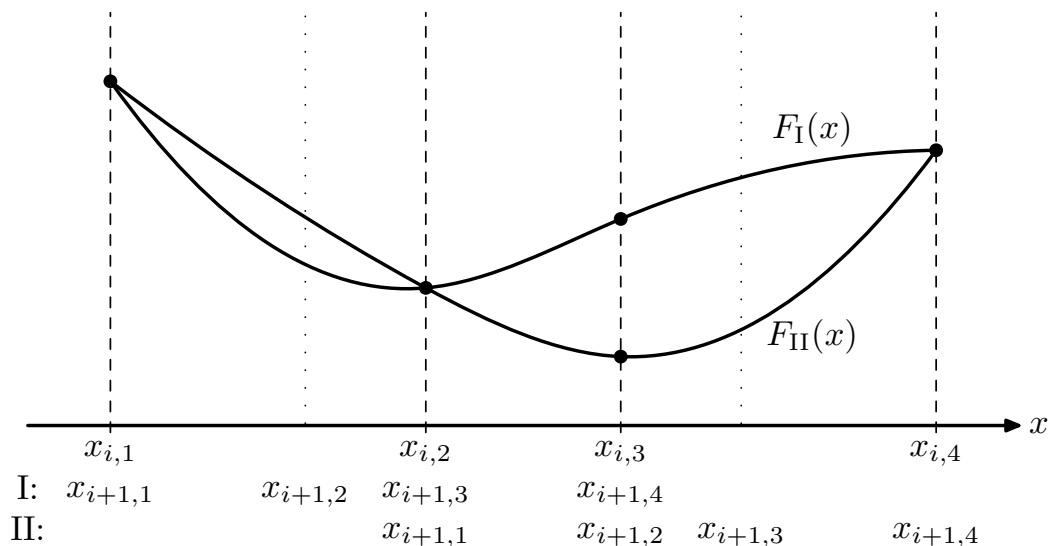
Celý algoritmu potom vyzerá takto:

1. Vyber počiatočný interval $\langle a_0, b_0 \rangle$ interval $\langle a, b \rangle$.
2. Rozdeľ interval $\langle a_i, b_i \rangle$ podľa *zlatého rezu*, t.j.

$$\begin{aligned} x_{i,1} &= a_i, \\ x_{i,2} &= a_i + (1 - \varphi)(b_i - a_i), \\ x_{i,3} &= a_i + \varphi(b_i - a_i), \\ x_{i,4} &= b_i. \end{aligned}$$

3. Ak $i = 0$, vypočítaj funkčné hodnoty funkcie $F(x)$ vo všetkých bodoch $x_{i,1}$ až $x_{i,4}$, inak vypočítaj funkčnú hodnotu funkcie $F(x)$ v novom bode ($x_{i,2}$ alebo $x_{i,3}$).
4. Ak $F(x_{i,2}) < F(x_{i,3})$, tak $\langle a_{i+1}, b_{i+1} \rangle = \langle x_{i,1}, x_{i,3} \rangle$, inak $\langle a_{i+1}, b_{i+1} \rangle = \langle x_{i,2}, x_{i,4} \rangle$.
5. Vráť sa k bodu 2. a proces opakuj.

Na obrázku 2.2 vidíme jeden krok metódy zlatého rezu. Už z predchádzajúceho kroku poznáme funkčné hodnoty funkcie $F(x)$ v bodoch $x_{i,1}$, $x_{i,2}$ a $x_{i,4}$. Vypočítame funkčnú hodnotu v bode $x_{i,3}$. Táto hodnota môže byť nad alebo pod hodnotou $F(x_{i,2})$.⁵ Obe situácie sú v obrázku zobrazené. Funkcia $F_I(x)$ zobrazuje príklad funkcie, ktorá by prechádzala hornou hodnotou v bode $x_{i,3}$ a funkcia $F_{II}(x)$ zobrazuje príklad funkcie, ktorá by prechádzala dolnou hodnotou v bode $x_{i,3}$. Rovnako bodkovanou čiarkou sú označené možno body nového delenia podintervalu v ďalšom kroku. Pod osou x sú uvedené nové delenie v prípadoch funkcií $F_I(x)$ a $F_{II}(x)$. Na obrázku je takisto vidno, že minimá funkcií $F_I(x)$ a $F_{II}(x)$ sa nachádzajú na vybraných podintervaloch.



Obr. 2.2: Jeden krok metódy zlatého rezu.

2.3.2 Optimalizácia vo viacrozmernom priestore

Okrem optimalizácie na jednorozmernom priestore sa stretáva i s optimalizáciou na viacrozmernom priestore. Úloha je potom iba preformulovaná do viacrozmerného priestoru. Máme danú reálnu konečnú funkciu $F(x_1, x_2, \dots, x_m)$ na otvorenej množine $M \in \mathbb{R}^m$

$$F(x_1, x_2, \dots, x_m) : M \rightarrow \mathbb{R}.$$

⁵Ak by nastal prípad, že $F(x_{i,2}) = F(x_{i,3})$, tak z predpokladov funkcie vyplýva, že minimum je na intervale $\langle x_{i,2}, x_{i,3} \rangle$. Vtedy hociktorá voľba podintervalu nám minimum zahŕňa.

Úlohou je nájsť extrém tejto funkcie na množine M . Je prirodzené v tomto prípade definovať m -ticu premenných ako vektor \mathbf{x}

$$\mathbf{x} := \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{pmatrix}.$$

V našej práci sa budeme zaoberať *Newtonovej metóde* a *kvázi-Newtonové metódy*.

2.3.2.1 Newtonova metóda

Newtonova metóda použitá pri optimalizácii vo viacrozmernom priestore je prirodzené rozšírenie z jednorozmerného prípadu. Základom je aproximácia funkcie $F(\mathbf{x})$ Taylorovým radom 2. stupňa $T_2^{F,\mathbf{a}}(\mathbf{x})$. Stupeň Taylorovho radu je najmenší, pri ktorom má aproximovaná funkcia extrém. Taylorov rad 2. stupňa funkcie $F(\mathbf{x})$ v m -rozmernom priestore v bode \mathbf{x}_i je

$$T_2^{F,\mathbf{x}_i}(\mathbf{x}) = F(\mathbf{x}_i) + \mathbf{D}F(\mathbf{x}_i)^T (\mathbf{x} - \mathbf{x}_i) + \frac{1}{2} (\mathbf{x} - \mathbf{x}_i)^T \mathbf{D}^2 F(\mathbf{x}_i) (\mathbf{x} - \mathbf{x}_i), \quad (2.33)$$

kde $\mathbf{D}F(\mathbf{x}_i)$ predstavuje vektor prvých derivácií funkcie $F(\mathbf{x})$ (*gradient*) v bode \mathbf{x}_i

$$\mathbf{D}F(\mathbf{x}_i) := \begin{pmatrix} \frac{\partial F}{\partial x_1} \\ \frac{\partial F}{\partial x_2} \\ \vdots \\ \frac{\partial F}{\partial x_m} \end{pmatrix} (\mathbf{x}_i),$$

a $\mathbf{D}^2 F(\mathbf{x}_i)$ maticu druhých derivácií funkcie $F(\mathbf{x})$ v bode \mathbf{x}_i (*Hessova matica*)

$$\mathbf{D}^2 F(\mathbf{x}_i) := \begin{pmatrix} \frac{\partial^2 F}{\partial x_1^2} & \frac{\partial^2 F}{\partial x_1 \partial x_2} & \cdots & \frac{\partial^2 F}{\partial x_1 \partial x_m} \\ \frac{\partial^2 F}{\partial x_2 \partial x_1} & \frac{\partial^2 F}{\partial x_2^2} & \cdots & \frac{\partial^2 F}{\partial x_2 \partial x_m} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 F}{\partial x_m \partial x_1} & \frac{\partial^2 F}{\partial x_m \partial x_2} & \cdots & \frac{\partial^2 F}{\partial x_m^2} \end{pmatrix} (\mathbf{x}_i).$$

Ak predpokladáme, že funkcia $F(\mathbf{x})$ je triedy $\mathcal{C}^2(M)$, tak nutnou podmienkou pre extrém je nulový gradient funkcie $F(\mathbf{x})$

$$\mathbf{D}F(\mathbf{x}_m) = 0, \quad (2.34)$$

kde \mathbf{x}_m je bod, kde je extrém funkcie. Zisťujeme extrém Taylorovho radu 2. stupňa funkcie $F(\mathbf{x})$

$$\mathbf{D}T_2^{F,\mathbf{a}}(\mathbf{x}) = 0.$$

Riešenie má potom tvar

$$\mathbf{x}_{i+1} = \mathbf{x}_i - \left(\mathbf{D}^2 F(\mathbf{x}_i) \right)^{-1} \mathbf{D}F(\mathbf{x}_i), \quad (2.35)$$

ktoré existuje práve jedno, ak je matica $\mathbf{D}^2 F(\mathbf{x}_i)$ regulárna, inak existuje nekonečne veľa riešení (rozdiel počtu neznámych m a hodnosti matice $h(\mathbf{D}^2 F(\mathbf{x}_i))$ určuje dimenziu priestoru riešení).

2.3.2.2 Kvázi-Newtonove metódy

Od Newtonovej metódy sú odvodené iné kvázi-Newtonove metódy, ktoré fungujú rovnako s jedným rozdielom. Hessova matica v nich nie je rátaná priamo, ale je odhadovaná na základe gradientov, v každom kroku je aktualizovaná. Zatiaľ čo počet prvkov gradientu rastie lineárne s počtom parametrov, rastie počet prvkov Hessovej matice kvadraticky s počtom parametrov.

Aproximácia funkcie má tvar

$$\tilde{F}_i(\mathbf{x}) = F(\mathbf{x}_i) + \mathbf{D}F(\mathbf{x}_i)^T (\mathbf{x} - \mathbf{x}_i) + \frac{1}{2} (\mathbf{x} - \mathbf{x}_i)^T \mathbf{B}_i (\mathbf{x} - \mathbf{x}_i), \quad (2.36)$$

kde sme Hessovu maticu nahradili odhadom \mathbf{B}_i .

Keďže sa zmenilo iba označenie, tak môžeme rovno písať riešenie pre hľadanie extrémum pre túto funkciu ako

$$\tilde{\mathbf{x}}_{i+1} - \mathbf{x}_i = -(\mathbf{B}_i)^{-1} \mathbf{D}F(\mathbf{x}_i), \quad (2.37)$$

kde bod $\tilde{\mathbf{x}}_{i+1}$ neberieme ako nové \mathbf{x}_{i+1} , ale pravú stranu berieme ako smer, v ktorom hľadáme extrém.

Keď už máme daný smer hľadania extrémum, tak ide o problém optimalizácie v jednorozmernom priestore. Označme $\mathbf{p}_i := \tilde{\mathbf{x}}_{i+1} - \mathbf{x}_i$. Výsledok optimalizácie v jednorozmernom priestore označme

$$\mathbf{x}_{i+1} = \mathbf{x}_i + \alpha_i \mathbf{p}_i, \quad (2.38)$$

kde α_i je bezrozmerný parameter určujúci polohu bodu \mathbf{x}_{i+1} na priamke $\mathbf{x}_i + \alpha \mathbf{p}_i$.

Keď však budeme mať novú aproximáciu funkcie v novom bode \mathbf{x}_{i+1} , potrebujeme novú maticu \mathbf{B}_{i+1} . Funkcia bude mať tvar

$$\tilde{F}_{i+1}(\mathbf{x}) = F(\mathbf{x}_{i+1}) + \mathbf{D}F(\mathbf{x}_{i+1})^T (\mathbf{x} - \mathbf{x}_{i+1}) + \frac{1}{2} (\mathbf{x} - \mathbf{x}_{i+1})^T \mathbf{B}_{i+1} (\mathbf{x} - \mathbf{x}_{i+1}). \quad (2.39)$$

Na novú funkciu je kladená prirodzená požiadavka. Chceme, aby gradient funkcie $\tilde{F}_{i+1}(\mathbf{x})$ sa rovnal vypočítanému gradientu funkcie $F(\mathbf{x})$ v bodoch \mathbf{x}_i a \mathbf{x}_{i+1} . Gradient funkcie $\tilde{F}_{i+1}(\mathbf{x})$ sa rovná

$$\mathbf{D}\tilde{F}_{i+1}(\mathbf{x}) = \mathbf{D}F(\mathbf{x}_{i+1}) + \mathbf{B}_{i+1} (\mathbf{x} - \mathbf{x}_{i+1}). \quad (2.40)$$

V bode \mathbf{x}_{i+1} je táto podmienka automaticky splnená. Aby sme splnili podmienku aj v bode \mathbf{x}_i , tak dostaneme maticovú rovnicu

$$\mathbf{B}_{i+1} (\mathbf{x}_{i+1} - \mathbf{x}_i) = \mathbf{D}F(\mathbf{x}_{i+1}) - \mathbf{D}F(\mathbf{x}_i). \quad (2.41)$$

Zavedením označenia $\mathbf{s}_i := \mathbf{x}_{i+1} - \mathbf{x}_i$ a $\mathbf{y}_i := \mathbf{D}F(\mathbf{x}_{i+1}) - \mathbf{D}F(\mathbf{x}_i)$ sa zjednoduší tvar rovnice na

$$\mathbf{B}_{i+1} \mathbf{s}_i = \mathbf{y}_i. \quad (2.42)$$

Ak funkcia $F(\mathbf{x})$ je triedy $\mathcal{C}^2(M)$ (má spojité druhé parciálne derivácie na množine M), tak podľa Schwarzovej vety je Hessova matica symetrická. Preto i matica \mathbf{B}_{i+1} sa predpokladá symetrická. Táto matica má $\frac{m(m+1)}{2}$ nezávislých hodnôt, zatiaľ čo rovnica (2.42) poskytuje iba m väzieb (podmienok, rovníc). Je

teda badateľné, že pre $m \geq 2$ je potrebné dodať ešte dodatočnú podmienku pre jednoznačné určenie matice \mathbf{B}_{i+1} .

Ďalej budeme požadovať, aby nová matica \mathbf{B}_{i+1} sa od starej \mathbf{B}_i líšila minimálne. Na priestore matíc máme definované viaceré normy, pomocou ktorých by sme minimalizovali normu rozdielu matíc \mathbf{B}_{i+1} a \mathbf{B}_i . Iná voľba normy určuje novú kvázi-Newtonovu metódu. Pri voľbe váženej Frobeniovej normy sa dostávame ku dvom metódam: *Davidonov-Fletcherov-Powellov algoritmus* (DFP) a *Broydenov-Fletcherov-Goldfarbov-Shannov algoritmus* (BFGS). Tieto metódy sú si istým spôsobom duálne. Zatiaľ čo podmienka na minimum váženej Frobeniovej normy na rozdiel matíc $\|\mathbf{B}_{i+1} - \mathbf{B}_i\|_F$ poskytuje rovnicu pre aktualizáciu matice \mathbf{B}_{i+1}^{-1} (DFP), tak podmienka na minimum váženej Frobeniovej normy na rozdiel matíc $\|\mathbf{B}_{i+1}^{-1} - \mathbf{B}_i^{-1}\|_F$ poskytuje rovnicu pre aktualizáciu matice \mathbf{B}_{i+1} (BFGS). Ukazuje sa, že efektívnejšia pri optimalizácii je metóda BFGS, preto si rozoberieme práve túto.

2.3.2.3 Broydenov-Fletcherov-Goldfarbov-Shannov algoritmus

Pri hľadaní matice \mathbf{B}_{i+1} , čo spĺňa danú podmienku minimalizácie váženej Frobeniovej normy $\|\mathbf{B}_{i+1}^{-1} - \mathbf{B}_i^{-1}\|_F$, dostaneme vzťah pre aktualizáciu matice a to

$$\mathbf{B}_{i+1} = \mathbf{B}_i - \frac{\mathbf{B}_i \mathbf{s}_i \mathbf{s}_i^T \mathbf{B}_i}{\mathbf{s}_i^T \mathbf{B}_i \mathbf{s}_i} + \frac{\mathbf{y}_i \mathbf{y}_i^T}{\mathbf{y}_i^T \mathbf{s}_i}. \quad (2.43)$$

Celý algoritmus potom vyzerá týmto spôsobom:

1. Vyber počiatkový bod \mathbf{x}_0 .
2. Vypočítaj odhad počiatkovej matice \mathbf{B}_0 v bode \mathbf{x}_0 .
3. Vypočítaj počiatkový gradient $\mathbf{D}F(\mathbf{x}_0)$ v bode \mathbf{x}_0 .
4. Vypočítaj smer hľadania \mathbf{p}_i podľa rovnice (2.37).
5. Nájdi minimum \mathbf{x}_{i+1} na priamke $\mathbf{x}_i + \alpha \mathbf{p}_i$.
6. Vypočítaj gradient $\mathbf{D}F(\mathbf{x}_{i+1})$ v bode \mathbf{x}_{i+1} .
7. Vypočítaj novú maticu \mathbf{B}_{i+1} podľa (2.43).
8. Vráť sa k bodu 4. a proces opakuj.

3. Výsledky, Systém He

3.1 Optimalizácia

V celej nasledujúcej kapitole nebudú vlnové funkcie v normovanom tvare. Keďže budeme používať Metropolisov-Hastingsov algoritmus, tak nás bude zaujímať pomer týchto funkcií (normovacie konštanty sa vykrátia). Rovnako používame atómové jednotky (energie sú v Hartree, vzdialenosti v Bohrových polomeroch atď.), teda pokiaľ nie sú uvedené explicitne, tak sa predpokladajú atómové.

Základná idea optimalizácie vyzerá nasledujúco. Základnou úlohou je vyrátať nasledujúci výraz

$$\frac{\langle \psi | \widehat{H} \psi \rangle}{\langle \psi | \psi \rangle}.$$

Pre Monte Carlo metódu volíme hustotu pravdepodobnosti ako $\varrho(\mathbf{r}) := \frac{\psi^2(\mathbf{r})}{\langle \psi | \psi \rangle}$ a vyčíslovanú funkciu ako $A(\mathbf{r}) := \frac{\widehat{H}\psi}{\psi}(\mathbf{r})$ (*lokálna energia*). Keďže v Metropolisovej-Hastingsovej metóde potrebujeme iba pomer hustôt $\varrho(\mathbf{r})$, nepotrebujeme rátať normalizačnú konštantu. Pomocou Metropolisovej-Hastingsovej metódy vygenerujeme vzorky \mathbf{r}_i , na ktorých vypočítame veličiny $A_i := A(\mathbf{r}_i)$. Pomocou štatistického spracovania a metódy blokovania vypočítame strednú hodnotu $\langle A \rangle$ s odhadom chyby (ak je chyby veľká, tak dopočítame ďalšie vzorky). Tak dostaneme pre dané parametre vlnovej funkcie jej energiu i odhad.

Tieto hodnoty používame v optimalizačných metódach, pre jeden parameter v metóde zlatého rezu, pre viac parametrov v BFGS metóde. Metóda BFGS vyžaduje znalosť výpočtu gradientu. Preto si najprv musíme vypočítať parciálnu deriváciu variačnej energie $E = \frac{\langle \psi(p) | \widehat{H} \psi(p) \rangle}{\langle \psi(p) | \psi(p) \rangle}$ podľa parametra

$$\begin{aligned} \frac{\partial}{\partial p} \left(\frac{\langle \psi(p) | \widehat{H} \psi(p) \rangle}{\langle \psi(p) | \psi(p) \rangle} \right) &= \frac{\frac{\partial}{\partial p} \langle \psi(p) | \widehat{H} \psi(p) \rangle - \frac{\langle \psi(p) | \widehat{H} \psi(p) \rangle}{\langle \psi(p) | \psi(p) \rangle} \frac{\partial}{\partial p} \langle \psi(p) | \psi(p) \rangle}{\langle \psi(p) | \psi(p) \rangle} = \\ &= \frac{\left\langle \psi^2(p) \left[\frac{\partial}{\partial p} \frac{\widehat{H}\psi(p)}{\psi(p)} + 2 \left(\frac{\widehat{H}\psi(p)}{\psi(p)} - E(p) \right) \frac{\partial}{\partial p} \frac{\psi(p)}{\psi(p)} \right] \right\rangle}{\langle \psi(p) | \psi(p) \rangle}. \end{aligned}$$

Vidíme, že stačí iba generovať vzorky podľa rovnakého rozdelenia, len budeme vyčíslovať inú funkciu $A(\mathbf{r})$.

V BFGS metóde je potrebné robiť optimalizáciu aj na jednorozmernom priestore. Tam využijeme metódu zlatého rezu.

Budeme sa zaoberať singletovým a tripletovým stavom atómu hélia v nerelativistickom priblížení a v Bornovej-Oppenheimerovej aproximácii. Vlnová funkcia teda závisí iba od premenných 2 elektrónov. Voľba spinového stavu nám uľahčuje voľbu vlnovej funkcie závislej iba od polohových premenných 2 elektrónov. Podľa 1.2.3 je vlnová funkcia závislá od priestorových súradníc 2 elektrónov pri singletovom stave symetrická a pri tripletovom stave antisymetrická.

Súradnice elektrónov označíme $\mathbf{r}_1 = (x_1, y_1, z_1)$ a $\mathbf{r}_2 = (x_2, y_2, z_2)$.

Podľa 1.2.4 platí, že sa môžeme obmedziť na reálne vlnové funkcie, lebo aj tie zahŕňajú všetky riešenia a neide o obmedzenie. Podľa 1.2.7 môžeme voliť funkciu závislej od vzdialeností častíc

$$\psi(\mathbf{r}_j) = \psi_f(r_1, r_2, r_{21}),$$

kde r_1, r_2 a r_{21} sú definované ako euklidovské normy vektorov $\mathbf{r}_1, \mathbf{r}_2$ a $\mathbf{r}_{21} := \mathbf{r}_2 - \mathbf{r}_1$ ¹

$$r_1(\mathbf{r}_1) := \|\mathbf{r}_1\|_2 = (x_1^2 + y_1^2 + z_1^2)^{1/2},$$

$$r_2(\mathbf{r}_2) := \|\mathbf{r}_2\|_2 = (x_2^2 + y_2^2 + z_2^2)^{1/2},$$

$$r_{21}(\mathbf{r}_1, \mathbf{r}_2) := \|\mathbf{r}_2 - \mathbf{r}_1\|_2 = \left[(x_2 - x_1)^2 + (y_2 - y_1)^2 + (z_2 - z_1)^2 \right]^{1/2}.$$

Ak by sme mali vlnovú funkciu zvolenú v tvare

$$\psi(\mathbf{r}_1, \mathbf{r}_2) = e^{f(r_1)} e^{f(r_2)} e^{g(r_{21})}, \quad (3.1)$$

resp. v tvare

$$\psi(\mathbf{r}_1, \mathbf{r}_2) = F(r_1)F(r_2)G(r_{21}). \quad (3.2)$$

Hamiltonián má tvar (v atómových jednotkách, v súradnicovej reprezentácii)

$$\widehat{H} = -\frac{1}{2} \left(\frac{\partial^2}{\partial x_1^2} + \frac{\partial^2}{\partial y_1^2} + \frac{\partial^2}{\partial z_1^2} \right) - \frac{1}{2} \left(\frac{\partial^2}{\partial x_2^2} + \frac{\partial^2}{\partial y_2^2} + \frac{\partial^2}{\partial z_2^2} \right) - \frac{2}{r_1} - \frac{2}{r_2} + \frac{1}{r_{21}}. \quad (3.3)$$

Pri výpočtoch sa nám hodia nasledujúce parciálne derivácie

$$\begin{aligned} \frac{\partial r_i}{\partial a_i} &= \frac{a_i}{(x_i^2 + y_i^2 + z_i^2)^{1/2}} = \frac{a_i}{r_i}, \\ \frac{\partial r_{21}}{\partial a_1} &= \frac{a_1 - a_2}{[(x_2 - x_1)^2 + (y_2 - y_1)^2 + (z_2 - z_1)^2]^{1/2}} = \frac{a_1 - a_2}{r_{21}}, \\ \frac{\partial r_{21}}{\partial a_2} &= \frac{a_2 - a_1}{[(x_2 - x_1)^2 + (y_2 - y_1)^2 + (z_2 - z_1)^2]^{1/2}} = \frac{a_2 - a_1}{r_{21}}, \\ \frac{\partial}{\partial a_i} \left(\frac{1}{r_i} \right) &= -\frac{a_i}{(x_i^2 + y_i^2 + z_i^2)^{3/2}} = -\frac{a_i}{r_i^3}, \\ \frac{\partial}{\partial a_1} \left(\frac{1}{r_{21}} \right) &= \frac{a_2 - a_1}{[(x_2 - x_1)^2 + (y_2 - y_1)^2 + (z_2 - z_1)^2]^{3/2}} = \frac{a_2 - a_1}{r_{21}^3}, \\ \frac{\partial}{\partial a_2} \left(\frac{1}{r_{21}} \right) &= \frac{a_1 - a_2}{[(x_2 - x_1)^2 + (y_2 - y_1)^2 + (z_2 - z_1)^2]^{3/2}} = \frac{a_1 - a_2}{r_{21}^3}, \end{aligned}$$

kde $a \in \{x, y, z\}$ a $i \in \{1, 2\}$.

¹Velkosť vektora môže byť označená r_{21} i r_{12} .

Vypočítali sme príslušné parciálne derivácie $\psi(\mathbf{r}_1, \mathbf{r}_2)$

$$\begin{aligned}
\frac{\partial \psi}{\partial a_1}(\mathbf{r}_1, \mathbf{r}_2) &= \left(f'(r_1) \frac{\partial r_1}{\partial a_1} + g'(r_{21}) \frac{\partial r_{21}}{\partial a_1} \right) \psi(\mathbf{r}_1, \mathbf{r}_2) = \\
&= \left(f'(r_1) \frac{a_1}{r_1} + g'(r_{21}) \frac{a_1 - a_2}{r_{21}} \right) \psi(\mathbf{r}_1, \mathbf{r}_2), \\
\frac{\partial \psi}{\partial a_2}(\mathbf{r}_1, \mathbf{r}_2) &= \left(f'(r_2) \frac{\partial r_2}{\partial a_2} + g'(r_{21}) \frac{\partial r_{21}}{\partial a_2} \right) \psi(\mathbf{r}_1, \mathbf{r}_2) = \\
&= \left(f'(r_2) \frac{a_2}{r_2} + g'(r_{21}) \frac{a_2 - a_1}{r_{21}} \right) \psi(\mathbf{r}_1, \mathbf{r}_2), \\
\frac{\partial^2 \psi}{\partial a_1^2}(\mathbf{r}_1, \mathbf{r}_2) &= \left[\left(f'(r_1) \frac{a_1}{r_1} + g'(r_{21}) \frac{a_1 - a_2}{r_{21}} \right)^2 + \right. \\
&\quad + \left(f''(r_1) \left(\frac{a_1}{r_1} \right)^2 + f'(r_1) \frac{1}{r_1} - f'(r_1) \frac{a_1^2}{r_1^3} \right) + \\
&\quad \left. + \left(g''(r_{21}) \left(\frac{a_1 - a_2}{r_{21}} \right)^2 + g'(r_{21}) \frac{1}{r_{21}} - g'(r_{21}) \frac{(a_1 - a_2)^2}{r_{21}^3} \right) \right] \psi(\mathbf{r}_1, \mathbf{r}_2), \\
\frac{\partial^2 \psi}{\partial a_2^2}(\mathbf{r}_1, \mathbf{r}_2) &= \left[\left(f'(r_2) \frac{a_2}{r_2} + g'(r_{21}) \frac{a_2 - a_1}{r_{21}} \right)^2 + \right. \\
&\quad + \left(f''(r_2) \left(\frac{a_2}{r_2} \right)^2 + f'(r_2) \frac{1}{r_2} - f'(r_2) \frac{a_2^2}{r_2^3} \right) + \\
&\quad \left. + \left(g''(r_{21}) \left(\frac{a_2 - a_1}{r_{21}} \right)^2 + g'(r_{21}) \frac{1}{r_{21}} - g'(r_{21}) \frac{(a_2 - a_1)^2}{r_{21}^3} \right) \right] \psi(\mathbf{r}_1, \mathbf{r}_2).
\end{aligned}$$

Dosadením vzťahov predchádzajúcich derivácií do 3.3 dostaneme pre $\frac{\widehat{H}\psi}{\psi}(\mathbf{r}_1, \mathbf{r}_2)$ tvar

$$\begin{aligned}
\frac{\widehat{H}\psi}{\psi}(\mathbf{r}_1, \mathbf{r}_2) &= -\frac{1}{2}f'(r_1)^2 - \frac{1}{2}f''(r_1) - f'(r_1)\frac{1}{r_1} - \frac{2}{r_1} - \\
&\quad - \frac{1}{2}f'(r_2)^2 - \frac{1}{2}f''(r_2) - f'(r_2)\frac{1}{r_2} - \frac{2}{r_2} - \\
&\quad - g'(r_{21})^2 - g''(r_{21}) - g'(r_{21})\frac{2}{r_{21}} + \frac{1}{r_{21}} + \\
&\quad + f'(r_1)g'(r_{21})\frac{(\mathbf{r}_1, \mathbf{r}_{21})}{r_1 r_{21}} - f'(r_2)g'(r_{21})\frac{(\mathbf{r}_2, \mathbf{r}_{21})}{r_2 r_{21}},
\end{aligned}$$

kde (\cdot, \cdot) je štandardný skalárny súčin. Postupujúc rovnako ako v prípade 1.2.7 dostaneme pre funkcie $f(x)$ a $g(x)$ podmienky

$$\begin{aligned}
f'(0) &= -2, \\
g'(0) &= \frac{1}{2},
\end{aligned}$$

respektíve ekvivalentné definované pre funkcie $F(x)$ a $g(x)$

$$\begin{aligned}
\frac{F'(0)}{F(0)} &= -2, \\
\frac{G'(0)}{G(0)} &= \frac{1}{2}.
\end{aligned}$$

3.1.1 Voľba tvaru vlnovej funkcie

Kľúčovým krokom v prípade variačnej metódy je voľba tvaru vlnovej funkcie. Voľba musí byť samozrejme v súlade s vlastnosťami vlnovej funkcie vypísaných v odstavci 1.2. Niektoré vieme splniť priamo voľbou funkcie, iné sa dajú splniť iba vhodnou voľbou parametrov.

Inšpiráciou nám môžu byť presné riešenia vodíku podobných systémov. Základný stav systému jedného elektrónu a jadra s nábojom Z má charakter

$$\psi_{0,H}(\mathbf{r}) = e^{-Zr} . \quad (3.4)$$

Rovnaký tvar dostaneme podľa 1.2.7 v okolí jadra. Ak by sme teda mali 2 nezávislé elektróny (nepôsobili by na seba Coulombickou interakciou), potom by riešenie spoločnej vlnovej funkcie sme dostali zo súčiny samotných vlnových funkcií elektrónov

$$\psi_{nez}(\mathbf{r}_1, \mathbf{r}_2) = e^{-Zr_1} e^{-Zr_2} . \quad (3.5)$$

Zaujímáť nás však bude člen funkcie, ktorý bude závisieť od r_{21} , ktorý bude vkladať koreláciu pohybu elektrónov. Z výpočtových dôvodov je výhodnejšie voliť túto funkciu v exponenciálnej funkcii

$$\psi_1(\mathbf{r}_1, \mathbf{r}_2) = e^{-Zr_1 - Zr_2 + g(r_{21})} . \quad (3.6)$$

Tento člen musí pre veľké vzdialenosti r_{21} byť nezávislý od r_{21} . Preto ho volíme v tvare

$$g(x) = \frac{P^n(x)}{Q^n(x)} ,$$

kde $P^n(x)$ a $Q^n(x)$ sú polynómy stupňa n . Ak majú oba polynómy rovnaký stupeň, tak limitne pre $x \rightarrow \infty$ sa funkcia $g(x)$ správa ako konštanta, čo sme požadovali. Podľa 1.2.7 viem však aj, ako má funkcia $g(x)$ limitne vyzeráť pre $x \rightarrow 0$. Musíme dostať limitne závislosť $\frac{x}{2}$.

Pre vlnovú funkciu s jedným parametrom a sme vybrali preto tvar

$$\psi_{1P}(\mathbf{r}_1, \mathbf{r}_2) = e^{-2r_1 - 2r_2 + \frac{r_{21}}{2(1+ar_{21})}} . \quad (3.7)$$

Pre vlnovú funkciu s viacerými premennými sme vybrali tvar, kde vo funkcii $g(x)$ sme vybrali kvadratické polynómy, rovnako sme zamenili náboj pri r_1 a r_2 parametrami, pridali sme aj člen, ktorý zodpovedá orbitálom typu p a funkciu sme ešte ošetrili vhodnou symetriou.

$$\psi_{7P}(\mathbf{r}_1, \mathbf{r}_2) = (1 \pm \hat{P}_{12}) \left[(1 + ar_1) e^{br_1 + cr_2 + \frac{d+er_{21}}{f^2+(r_{21}-g)^2}} \right] . \quad (3.8)$$

3.1.2 Výber generovania vzoriek

Spôsobov ako generovať vzorky do Metropolisovho-Hastingsovho algoritmu je mnoho. Preto sme si na jednoduchšej vlnovej funkcii ψ_{ne} z vyskúšali 4 rôzne generovania a vybrali najvhodnejšie.

Budeme generovať 3 súradnice dvoch elektrónov. Dokopy 6 súradníc, ktoré označíme $\mathbf{r} = (\mathbf{r}_1, \mathbf{r}_2)$. Použili sme nasledujúce 4 generovania:

1. ku polohe \mathbf{r}_i pričítame v každej zložke náhodné číslo z uniformného rozdelenia $\mathcal{U}(-R, R)$ (generovanie A)

$$\mathbf{r}'_i = \mathbf{r}_i + \mathbf{\Delta}, \quad \mathbf{\Delta} \in \mathcal{U}(-R, R)^6,$$

2. ku polohe \mathbf{r}_i pričítame v každej zložke náhodné číslo z normálneho rozdelenia $\mathcal{N}(0, R^2)$ (generovanie B)

$$\mathbf{r}'_i = \mathbf{r}_i + \mathbf{\Delta}, \quad \mathbf{\Delta} \in \mathcal{N}(0, R^2)^6,$$

3. každú polohu \mathbf{r}'_i vyberáme po zložkách z normálneho rozdelenia $\mathcal{N}(0, R^2)$ (generovanie C)

$$\mathbf{r}'_i \in \mathcal{N}(0, R^2)^6,$$

4. každú polohu \mathbf{r}'_i vyberáme s hustotou pravdepodobnosti (generovanie D)

$$\varrho(\mathbf{r}'_1, \mathbf{r}'_2) = \frac{\alpha e^{-\alpha r'_1}}{4\pi r_1'^2} \frac{\alpha e^{-\alpha r'_2}}{4\pi r_2'^2}.$$

Najprv je potrebné pri jednotlivých generovaniach vybrať vhodnú hodnotu parametra (R alebo α). Od tejto voľby závisí pravdepodobnosť prijatia p novej vzorky v Metropolisovom-Hastingsovom algoritme. O tejto pravdepodobnosti vieme povedať, že nechceme, aby bola nízka (blízko nule) lebo by sme často novú polohu zamietli, vzorky by boli silno korelované a potrebovali by sme veľa výpočtového času na dosiahnutie určitej presnosti. Vybrané generovanie má mať vlastnosť, že za pomerne krátky čas vybrané vzorky dokážu charakterizovať celé rozdelenie. To je problém generovaní A a B, lebo sú závislé od predchádzajúcej polohy. Ak by bolo R zvolené veľmi malé, tak získané vzorky sa nebudú nachádzať ďaleko od počiatočnej polohy, vzorky budú silno korelované a vzorky nebudú pokrývať celé rozdelenie iba okolie počiatočného bodu.

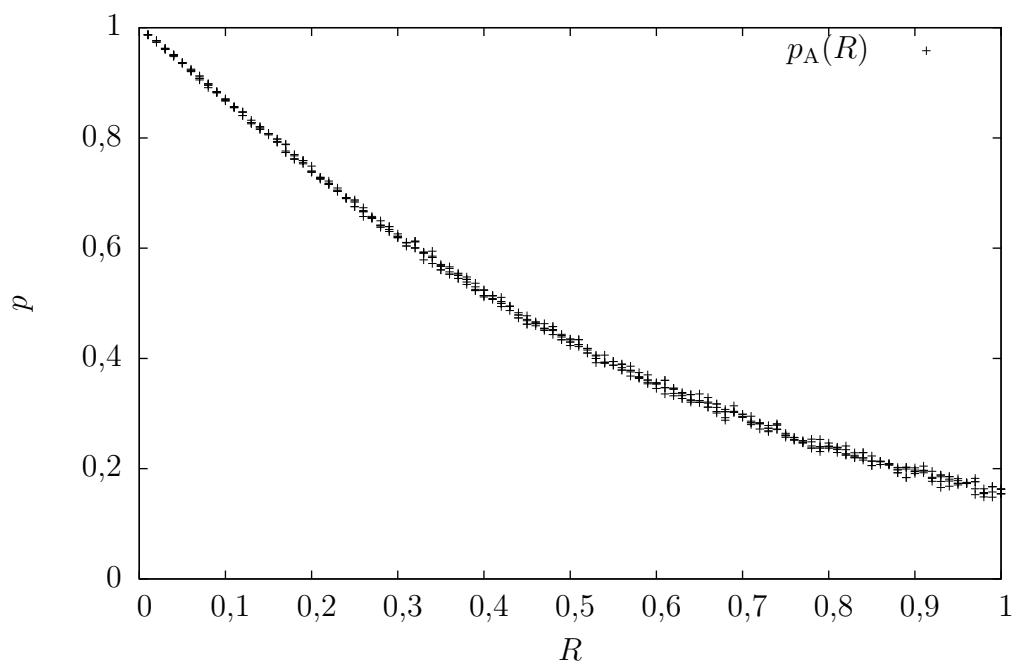
V grafoch 3.1 až 3.4 sme zobrazili závislosť pravdepodobnosti prijatia novej navrhutej polohy od parametra generovania na základe ktorých sme zvolili hodnotu parametra, ktorú sme ďalej používali. V každom grafe je pre jednu danú hodnotu parametra zobrazených 5 vypočítaných pravdepodobností.

V prípade generovania A (3.1) je badať, že pre veľké hodnoty parametra R pravdepodobnosť klesá k nule. Pre malé hodnoty parametra sú zase vybrané vzorky korelované. Preto je vhodné vybrať hodnotu parametra pri ktorom je pravdepodobnosť približne $p_A(R_A) \approx 0,5$. Preto sme vybrali hodnotu $R_A = 0,4$.

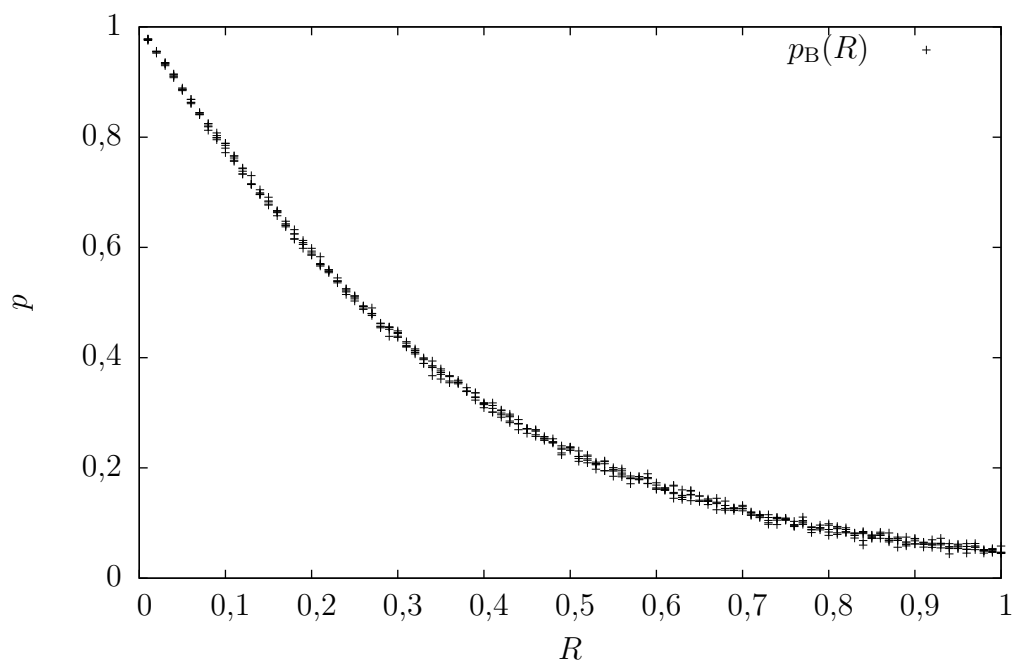
Generovanie B (3.2) má podobný problém ako generovanie A a podobný charakter grafu. Postupovali sme obdobne a vybrali hodnotu $R_B = 0,25$.

V prípade generovania C (3.3) vidíme, že rovnako pre malé i pre veľké hodnoty parametra klesá pravdepodobnosť ku nule. V oblasti od 0,2 až 0,5 vidíme, že dané hodnoty pravdepodobnosti sa výraznejšie líšia pre jednu hodnotu (väčší rozptyl hodnôt), preto sme radšej vybrali hodnotu parametra, kde je pravdepodobnosť nižšia, ale je malý rozptyl pravdepodobnosti $R_C = 0,6$.

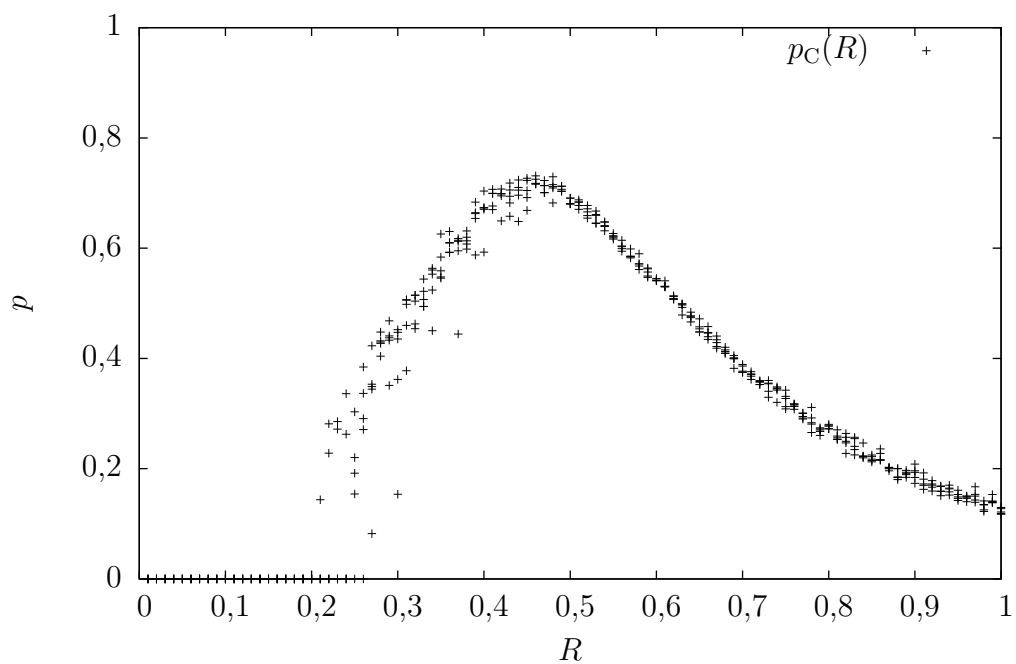
V poslednom prípade generovania (generovanie D 3.4) pravdepodobnosť je pre $\alpha = 0$ nulová a rovnako klesá k nule pre α rastúce do nekonečna. Hodnota parametra bola vybraná pri maximálnej pravdepodobnosti $\alpha_D = 1,4$.



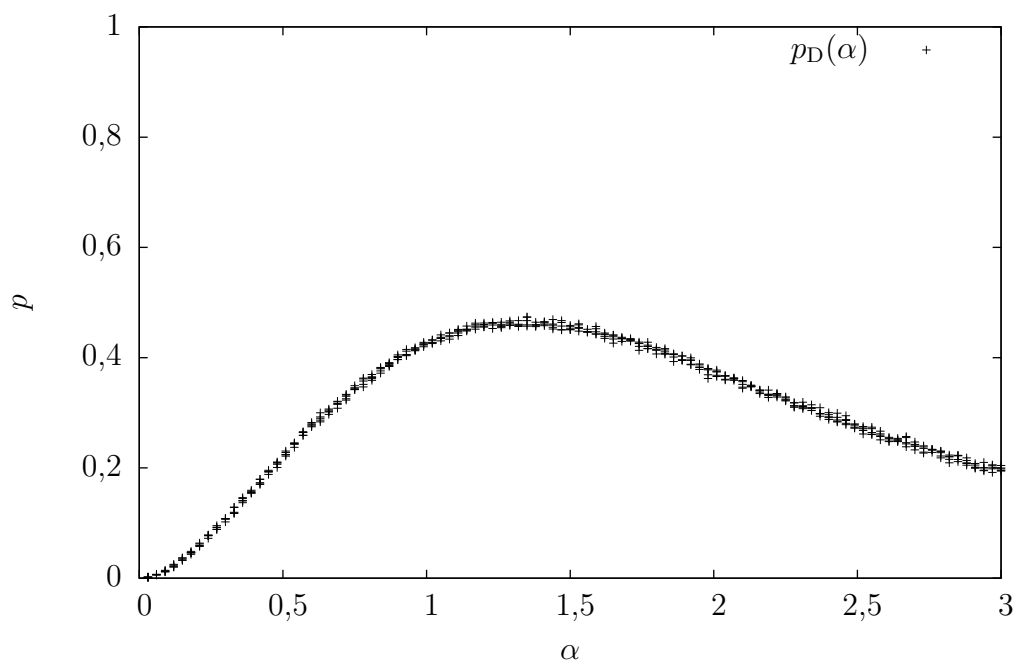
Obr. 3.1: Závislosť pravdepodobnosti výberu od parametra generovaného rozdelenia (generovanie A).



Obr. 3.2: Závislosť pravdepodobnosti výberu od parametra generovaného rozdelenia (generovanie B).



Obr. 3.3: Závislosť pravdepodobnosti výberu od parametra generovaného rozdelenia (generovanie C).



Obr. 3.4: Závislosť pravdepodobnosti výberu od parametra generovaného rozdelenia (generovanie D).

Pre vybrané hodnoty parametrov sme následne zisťovali koreláciu N vzoriek pri jednotlivých generovaniach. Na vybraných vzorkách \mathbf{r}_i sme vyčíslili hodnotu energie E_i a následne sme vypočítali odhad korelačnej funkcie $c(n)$ pre zopár prvých členov. Ten sme zvolili

$$c(n) := \frac{1}{N-n} \sum_{i=1}^{N-n} (E_i - \tilde{E})(E_{i+n} - \tilde{E}),$$

kde \tilde{E} je výberový priemer vzoriek E_i .

Aby sme mohli jednotlivé generovania porovnať, zobrazili sme normovaný odhad korelačnej funkcie $d(n)$

$$d(n) = \frac{c(n)}{c(0)}. \quad (3.9)$$

Uvedený odhad sme ráтали pre päť rôzne vygenerovaných N vzoriek.

V grafe 3.5 sme zobrazili pre každé generovanie pre päť rôznych množín vygenerovaných vzoriek závislosť normovaného odhadu korelačnej funkcie od relatívnej vzdialenosti vzoriek n . Samozrejme, od vybraného generovania požadujeme, aby korelácia bola najmenšia, lebo tá spôsobuje vyšší odhad chyby, čo zas vedie k dlhšiemu výpočtu na dosiahnutie danej chyby. Dá sa očakávať, že generovania, ktoré vyberajú na základe predchádzajúcej polohy, budú mať silnejšiu koreláciu. Toto môžeme na grafe 3.5 pozorovať pre generovania A a B. Rovnako je vidno, že pri generovaní C je niekedy korelácia silnejšia, niekedy slabšia, čo by mohlo byť spôsobené tým, že sa s malou šancou vygeneruje vzorka, ktorej pravdepodobnosť generovania je veľmi malá, čo v Metropolisovom-Hastingsovom algoritme spôsobí to, že sa je malá šanca, že sa ďalšie navrhované vzorky príjmu. Potom máme v množine vzoriek dlhú sekvenciu jednej hodnoty, čo spôsobuje vysokú koreláciu.

Z grafu 3.5 je samozrejme, že najslabšia korelácia bola v prípade generovania D. Preto sme si pre ďalšiu prácu vybrali práve toto generovanie.

3.1.3 Vlnová funkcia s jedným parametrom

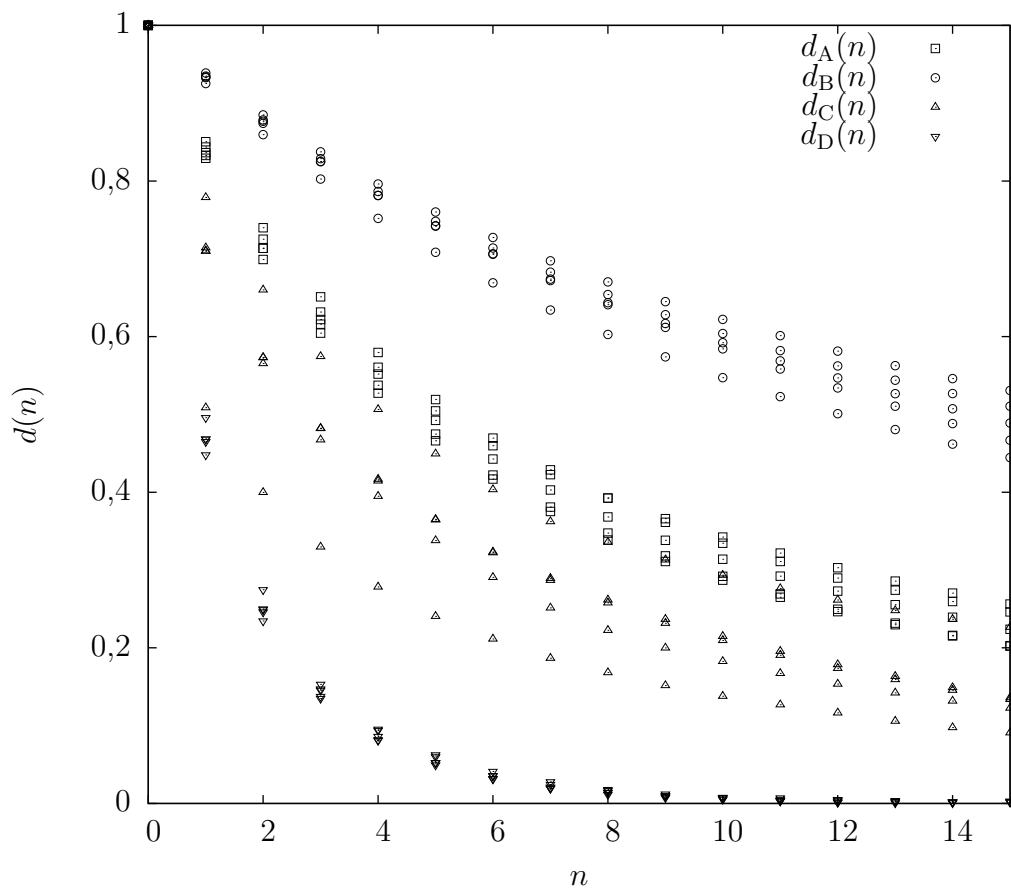
Ako vlnovú funkciu s jedným parametrom a sme zvolili vlnovú funkciu tvaru

$$\psi_{1P}(\mathbf{r}_1, \mathbf{r}_2) = e^{-2r_1 - 2r_2 + \frac{r_{21}}{2(1+ar_{21})}}. \quad (3.10)$$

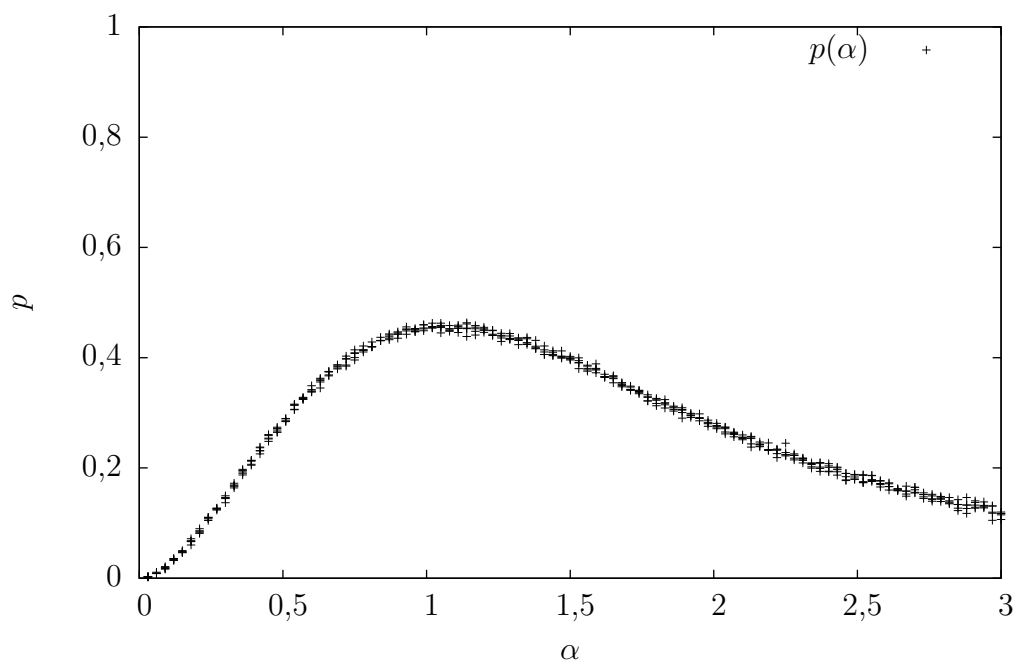
Najprv musíme určiť najvhodnejšiu hodnotu parametra generovania D. Pre nejaký parameter vlnovej funkcia (zvolili sme $a = 0$) sme zopakovali rovnaký proces, ako v prípade výberu generovania.

V grafe 3.6 sme zobrazili závislosť pravdepodobnosti prijatia novej navrhutej polohy od parametra generovania α (pre danú hodnotu parametra päťkrát). Na základe grafu sme vybrali hodnotu $\alpha_{1P} = 1, 1$.

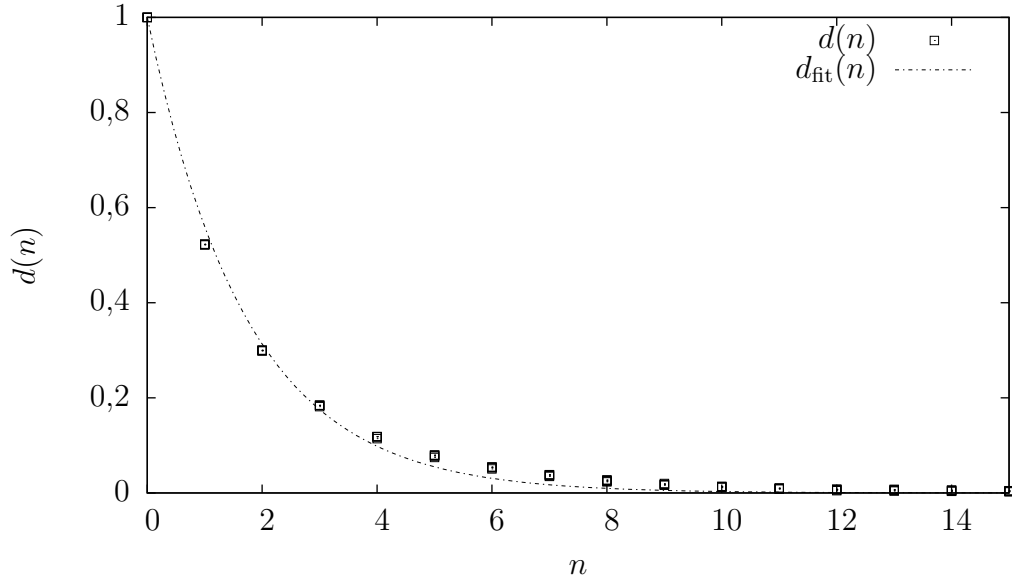
Rovnako sme v grafe 3.7 zobrazili závislosť normovaného odhadu korelačnej funkcie od relatívnej vzdialenosti vzoriek n (päťkrát pre každú hodnotu). Hodnotami sme preložili exponenciálnu funkciu $d_{\text{fit}}(n) = e^{-n/\tau}$, aby sme mohli porovnať, či má korelačná funkcia exponenciálny charakter. Preložená funkcia dosiahla parameter $\tau = (1, 72 \pm 0, 05)$, kde sme uviedli iba chybu spôsobenú prekladáním. Vzhľadom na graf si korelačná funkcia zachováva približne exponenciálny charakter.



Obr. 3.5: Závislosť normovaného odhadu korelačnej funkcie pre rôzne druhy generovania.



Obr. 3.6: Závislosť pravdepodobnosti výberu od parametra generovania v prípade vlnovej funkcie jedného parametra.



Obr. 3.7: Závislosť normovaného odhadu korelačnej funkcie.

Potom sme spustili samotný program optimalizácie. Použili sme metódu zlatého rezu na hľadanie extrému v prípade jedného parametra. Nakoľko nerátame funkčné hodnoty presne, upravili sme podmienky na rozhodnutie o delení. Rozhodnúť môžeme iba ak je rozdiel funkčných hodnôt väčší ako t -násobok súčtu chýb funkčných hodnôt, kde t sme zvolili $t = 2.0$. Ako počiatočný interval sme zvolili interval $\langle 0, 5 \rangle$.

V grafe 3.8 sme zobrazili závislosť vypočítanej energie $E(a)$ stavu od hodnoty parametra a s chybovými úsečkami. Rovnako sme bodmi preložili spline krivku $E_{\text{fit}}(a)$. Táto krivka nedodáva žiadnu ďalšiu informáciu, len vykresluje, ako by napríklad mohol vyzerat priebeh skutočnej funkcie $E(a)$.

Použitím jedného parametra sme sa dostali odhad energie základného stavu ako

$$E_{1P} = (-2, 8772 \pm 0, 0004), \quad (3.11)$$

pre parameter $a = 0, 172209$.

3.1.4 Vlnová funkcia s viacerými parametrami

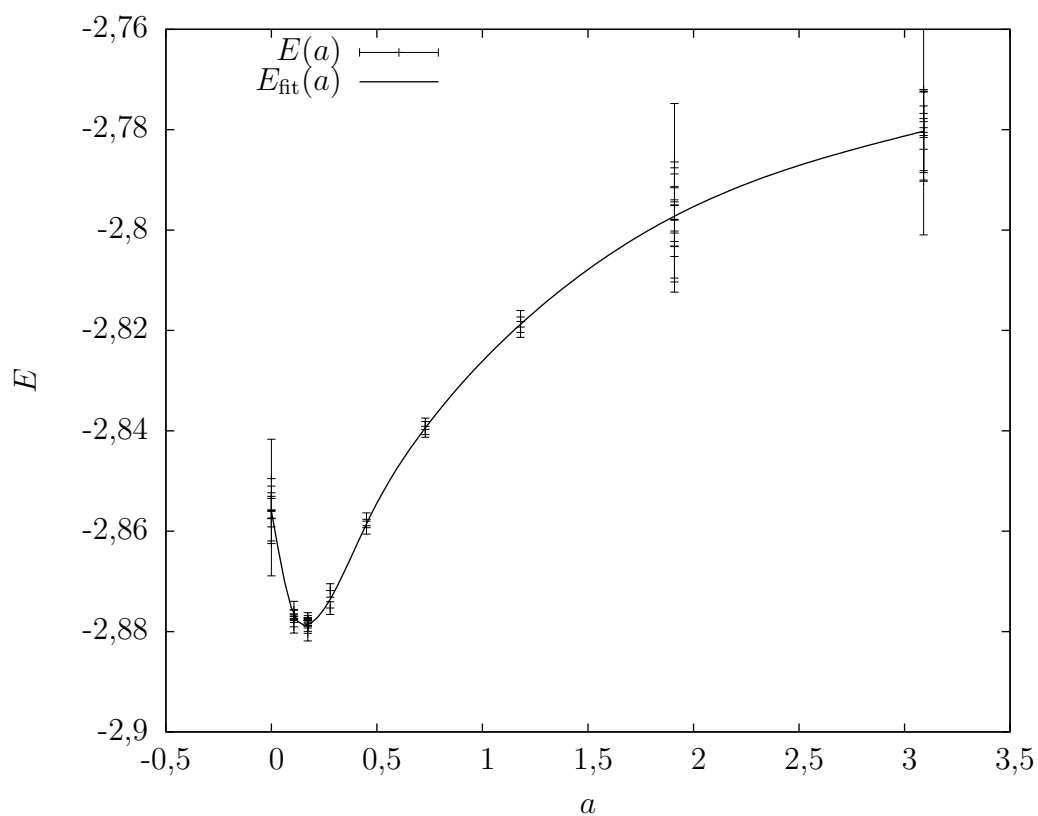
Ako vlnovú funkciu s viacerými parametrami sme zvolili funkciu tvaru

$$\psi_{7P}(\mathbf{r}_1, \mathbf{r}_2) = (1 \pm \hat{P}_{12}) \left[(1 + ar_1) e^{br_1 + cr_2 + \frac{d+er_1}{f^2+(r_{21}-g)^2}} \right].$$

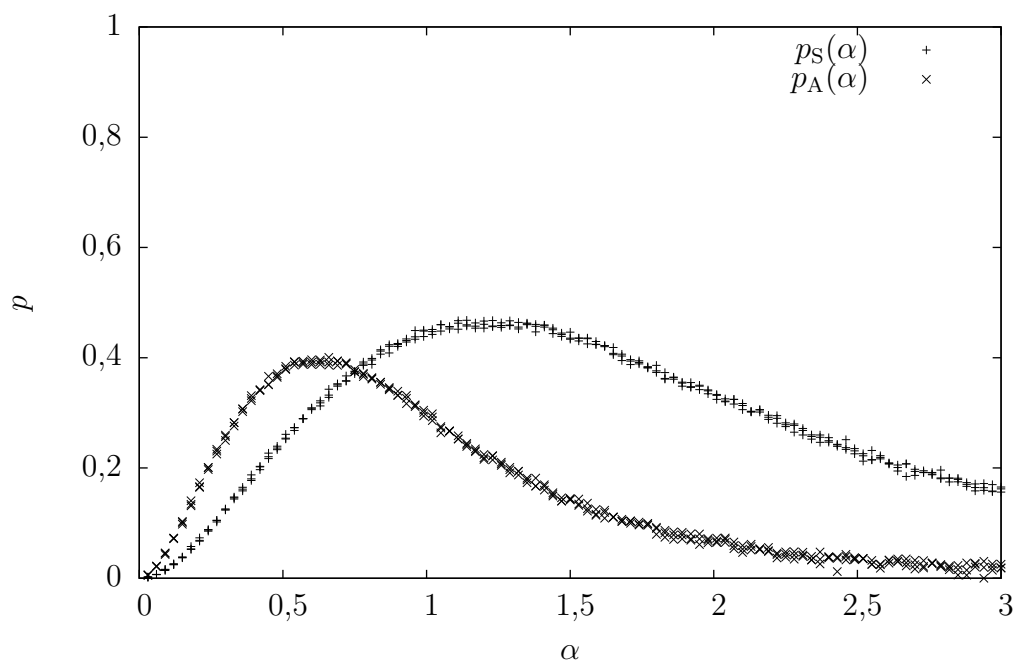
Počiatočné parametre sme zvolili buď podľa asymptotického správania alebo podľa hodnoty parametra a získanej v prípade vlnovej funkcie iba s jedným parametrom. Ich hodnoty sú uvedené v tabuľkách 3.1, 3.3 a 3.5.

Následne sme určili pre symetrickú i antisymetrickú funkciu najvhodnejší parameter generovania rovnako, ako v prechádzajúcich príkladoch.

V grafe 3.9 sme zobrazili pravdepodobnosť prijatia novej vygenerovanej vzorky od parametra generovania α pre symetrickú $p_S(\alpha)$ a antisymetrickú vlnovú funkciu $p_A(\alpha)$ (trikrát pre dané hodnoty α). Zvolili sme pre ďalšie výpočty hodnotu $\alpha_S = 1, 2$ pre symetrickú a $\alpha_A = 0, 6$ pre antisymetrickú vlnovú funkciu.

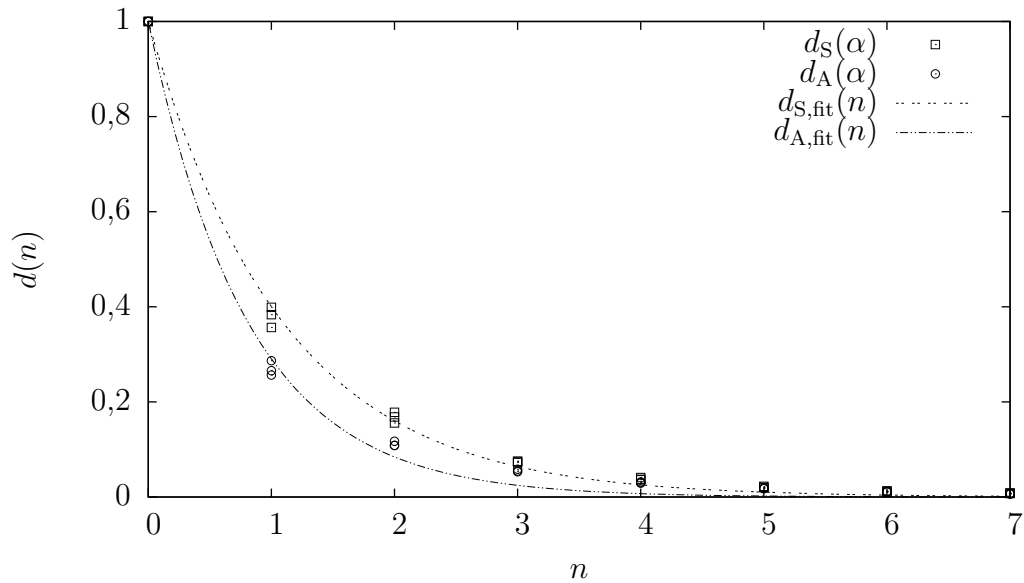


Obr. 3.8: Závislosť energie stavu od parametra vlnovej funkcie a .



Obr. 3.9: Závislosť pravdepodobnosti výberu od parametra generovania v prípade symetrickej a antisymetrickej vlnovej funkcie 7 parametrov.

Potom sme v grafe 3.10 zobrazili závislosť normovaného odhadu korelačnej funkcie od relatívnej vzdialenosti vzoriek n (trikrát pre každú hodnotu). Hodnotami sme preložili exponenciálnu funkciu $d_{X,\text{fit}}(n) = e^{-n/\tau_X}$ ($X \in \{S, A\}$). Preložená funkcia dosiahla parameter $\tau_S = (1,09 \pm 0,03)$ pre symetrickú vlnovú funkciu a $\tau_A = (0,81 \pm 0,04)$ pre antisymetrickú vlnovú funkciu (uvedená iba chyba spôsobená preložením). Normovaná korelačná funkcia pre oba prípady má približne exponenciálny charakter, ale je viditeľne odchýlená od skutočnej závislosti. Závislosť korelačnej funkcie nám pomôže určiť, na počet krokov od prechádzajúcej polohy, aby sme sa minimalizovali koreláciu, a tým odhad chyby, aby sme ušetrili výpočtový čas. Veľa krokov nám zasa predĺži výpočtový čas, lebo budeme potrebovať veľa operácií na získanie jednej vzorky. Zvolili sme preto 3 kroky.



Obr. 3.10: Závislosť normovaného odhadu korelačnej funkcie pre symetrickú a antisymetrickú funkciu.

Následne sme spustili optimalizáciu pre tri počiatočné polohy. Síce je volená funkcia symetrická/antisymetrická voči zámene elektrónov, tak parametre nie sú volené symetricky.

Zatiaľ čo v prípade symetrickej funkcie stačí zvoliť parametre b a c rovnaké a optimalizácia nájde minimum, tak v prípade antisymetrickej funkcie je potrebné zvoliť počiatočné parametre b a c rôzne. Preto sme robili 3 optimalizácie. Jednu pre symetrickú funkciu (S1) a 2 pre antisymetrickú funkciu (A1, A2).

V tabuľkách 3.1, 3.3 a 3.5 sme vypísali priebeh optimalizácie pre hodnoty parametrov a hodnoty energie E (s odhadom chyby ΔE) v závislosti od čísla kroku optimalizácie pre všetky prípady S1, A1 a A2. Je vidno, ktoré parametre sa optimalizujú a ktoré nie. V prípade S1 sa optimalizujú prvé tri parametre, no parametre d až g sa zásadne nemenia. V tomto prípade sa zdá pre danú presnosť táto časť vlnovej funkcie preparametrizovaná. V prípade A1 a A2 sa tieto parametre čiastočne menili.

V tabuľkách 3.2, 3.4 a 3.6 sme vypísali vývoj hodnôt energie p (s odhadom chyby), viriálový pomer p^2 (s odhadom chyby) a veličiny ΔH pre všetky prípady

²Viriálový pomer sme definovali ako pomer strednej hodnoty potenciálnej energie $\langle V \rangle$ a

S1, A1 a A2.

i	a	b	c	d	e	f	g	E	ΔE
0	0,100	-2,000	-2,000	-107,0	-17,00	1,000	-7,000	-2,86985	0,00050
1	0,123	-1,965	-1,972	-109,3	-15,99	0,900	-6,389	-2,88867	0,00050
2	0,140	-1,932	-1,954	-108,8	-16,17	0,931	-6,567	-2,89064	0,00050
3	0,183	-1,840	-1,946	-108,0	-16,77	0,926	-6,635	-2,89169	0,00050
4	0,160	-1,826	-2,077	-108,4	-16,33	0,898	-6,720	-2,89384	0,00050
5	0,207	-1,641	-2,240	-108,2	-16,85	0,871	-6,815	-2,90048	0,00050
6	0,207	-1,641	-2,240	-108,2	-16,85	0,871	-6,815	-2,90059	0,00050
7	0,238	-1,554	-2,257	-107,3	-17,19	0,878	-6,899	-2,90057	0,00050
8	0,238	-1,554	-2,257	-107,3	-17,19	0,878	-6,899	-2,90031	0,00050
9	0,224	-1,616	-2,175	-107,3	-17,15	0,907	-6,849	-2,90086	0,00050
10	0,208	-1,612	-2,263	-107,9	-16,85	0,876	-6,896	-2,90062	0,00050
11	0,219	-1,598	-2,222	-107,7	-17,16	0,880	-6,880	-2,90171	0,00050
12	0,224	-1,592	-2,232	-108,3	-17,16	0,845	-6,938	-2,90265	0,00050
13	0,209	-1,624	-2,202	-109,3	-17,27	0,792	-6,868	-2,90153	0,00050
14	0,212	-1,605	-2,227	-109,4	-17,60	0,781	-6,840	-2,90191	0,00049
15	0,204	-1,613	-2,225	-110,3	-18,03	0,745	-6,766	-2,90107	0,00050
16	0,213	-1,596	-2,215	-108,4	-17,72	0,840	-6,864	-2,90111	0,00049
17	0,222	-1,593	-2,209	-107,7	-17,68	0,848	-6,874	-2,90172	0,00050

Tabuľka 3.1: Číslo kroku i , hodnoty parametrov po i -tom kroku optimalizácie, hodnoty energie E s odhadom chyby ΔE pre prípad S1.

Pre lepšiu predstavu sme v grafoch 3.11, 3.12 zobrazili vývoj energie stavu aj s chybou. Je vidieť, že približne 5 krokov sa energia stavu rapídne znížila a následne už len oscilovala okolo jednej hodnoty. V prípadoch A1 a A2 sa vzhľadom na veľkosť chyby nedá jednoznačne povedať, v ktorých z týchto lokálnych minim sa nachádza globálne minimum.

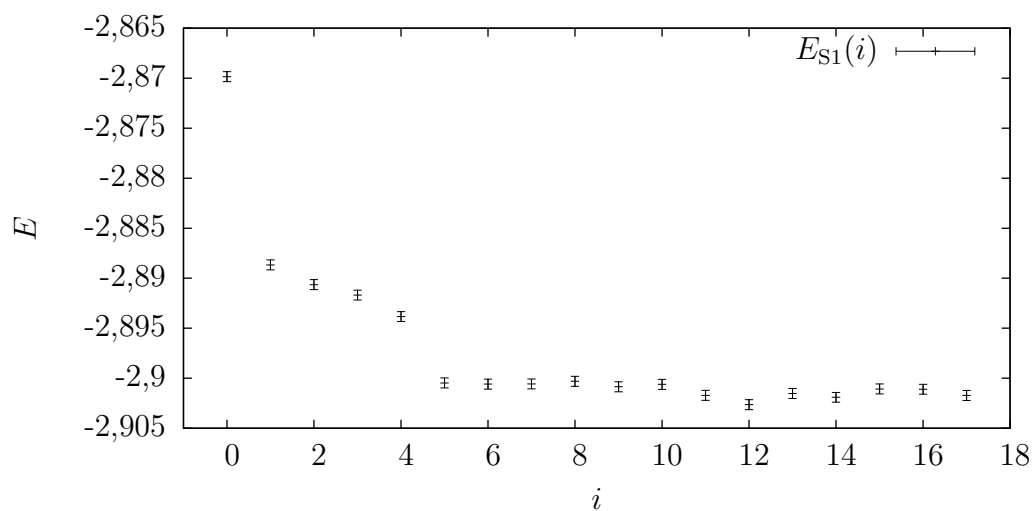
V grafe 3.13 sme zobrazili vývoj viriálového pomeru a správnu hodnotu pre presný stav -2 . Je vidieť, že v prvých piatich krokoch (keď sa znižovala energia) sa postupne hodnoty približovali ku správnej hodnote a následne sa hodnoty pohybovali okolo hodnoty -2 , no už nezmenšovali svoj rozptyl od správnej hodnoty.

V grafe 3.14 sme zobrazili vývoj veličiny ΔH . Obdobne ako pri iných veličinách i tu v prvých 5 krokoch je badať pokles veličiny ku správnej hodnote 0, ale následne sa už hodnota nezlepšuje.

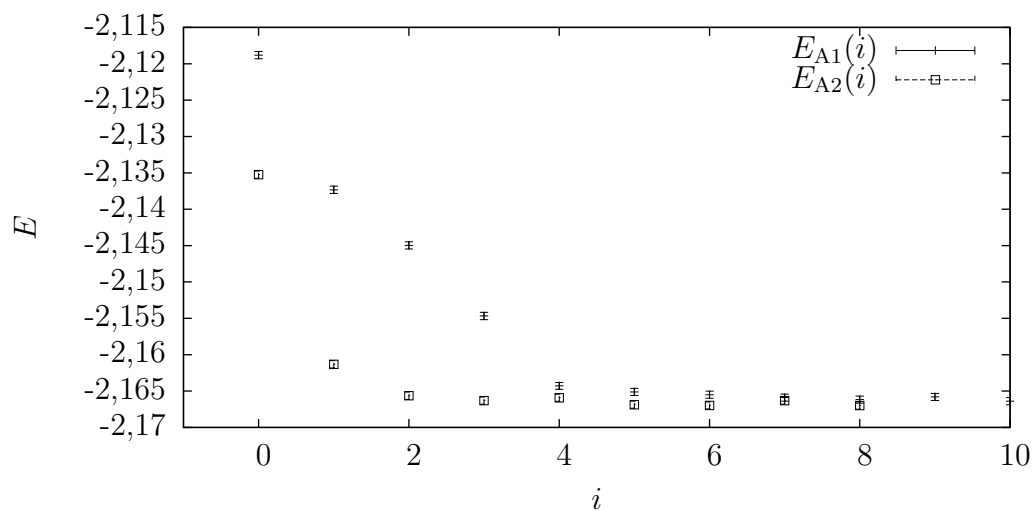
Na porovnanie najlepších vypočítaných hodnôt energie s experimentálnymi a vypočítanými hodnotami z iných zdrojov slúži tabuľka 3.7. Vidíme, že vypočítané hodnoty sú nižšie ako blízko-limitné hodnoty Hartreeovej-Fockovej metódy. Dokonca iba použitím jedného parametra. V prípade singletu sme použitím vlnovej funkcie so siedmymi parametrami znížili horný odhad energie základného stavu na rozdielom $E_{7P} - E_{\text{var}} = 0,0011$ od presného výpočtu (s presnosťou $\Delta E = 0,0005$). V prípade tripletu podobná vlnová funkcia so siedmymi parametrami

kinetickej energie $\langle T \rangle$

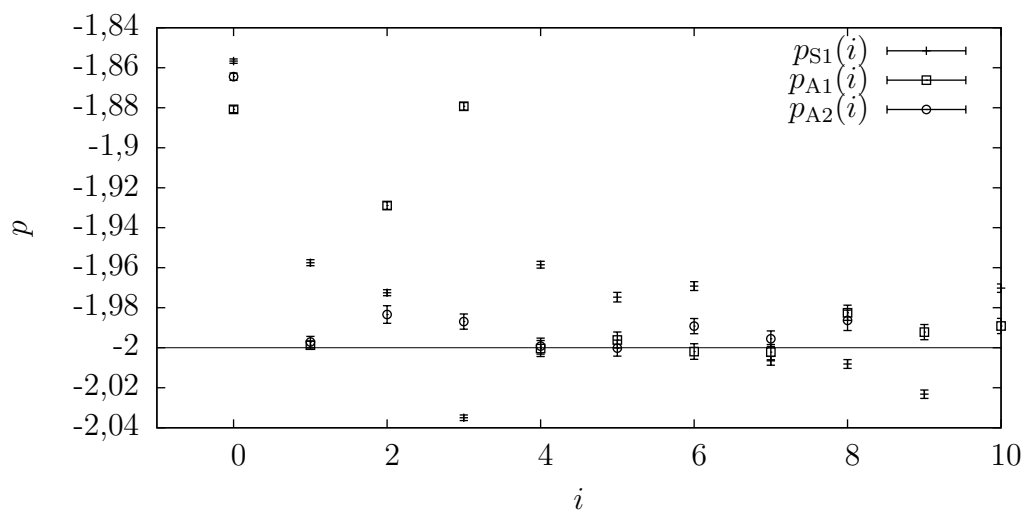
$$p := \frac{\langle V \rangle}{\langle T \rangle}. \quad (3.12)$$



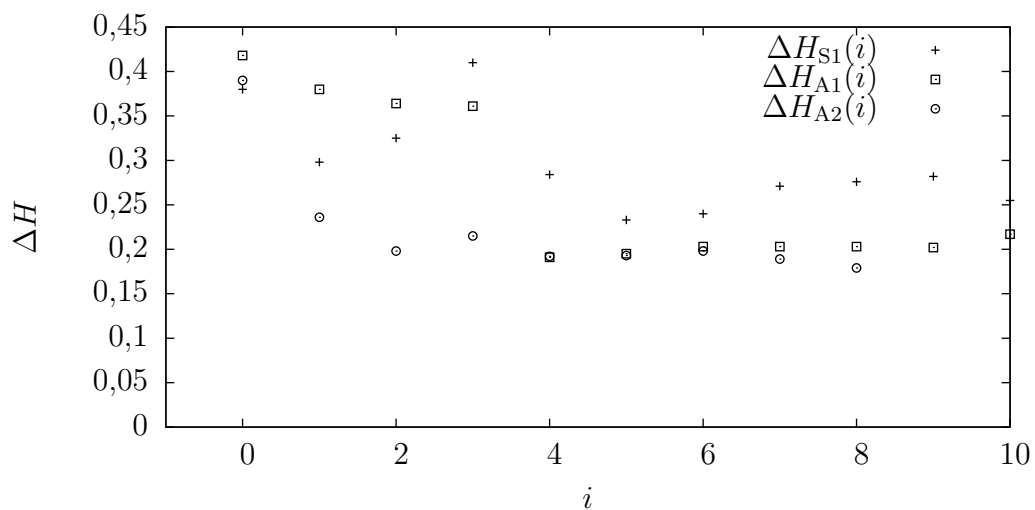
Obr. 3.11: Závislosť energie stavu E s odhadom chyby od čísla kroku optimalizácie i v prípade S1.



Obr. 3.12: Závislosť energie stavu E s odhadom chyby od čísla kroku optimalizácie i v prípade A1 a A2.



Obr. 3.13: Závislosť viriálového pomeru p s odhadom chyby od čísla kroku optimalizácie i v prípadoch S1, A1 a A2.



Obr. 3.14: Závislosť hodnoty ΔH od čísla kroku optimalizácie i v prípadoch S1, A1 a A2.

i	E	ΔE	p	Δp	ΔH
0	-2,86985	0,00050	-1,8567	0,0011	0,380
1	-2,88867	0,00050	-1,9575	0,0015	0,298
2	-2,89064	0,00050	-1,9725	0,0015	0,325
3	-2,89169	0,00050	-2,0350	0,0014	0,410
4	-2,89384	0,00050	-1,9585	0,0017	0,284
5	-2,90048	0,00050	-1,9747	0,0024	0,233
6	-2,90059	0,00050	-1,9692	0,0022	0,240
7	-2,90057	0,00050	-2,0065	0,0022	0,271
8	-2,90031	0,00050	-2,0081	0,0022	0,276
9	-2,90086	0,00050	-2,0232	0,0021	0,282
10	-2,90062	0,00050	-1,9702	0,0021	0,255
11	-2,90171	0,00050	-2,0023	0,0023	0,258
12	-2,90265	0,00050	-1,9934	0,0023	0,257
13	-2,90153	0,00050	-1,9961	0,0022	0,251
14	-2,90191	0,00049	-1,9921	0,0023	0,239
15	-2,90107	0,00050	-1,9935	0,0024	0,245
16	-2,90111	0,00049	-2,0054	0,0021	0,272
17	-2,90172	0,00050	-2,0049	0,0022	0,273

Tabuľka 3.2: Číslo kroku i , hodnoty energie E s odhadom chyby ΔE , viriálový pomer p s odhadom chyby Δp a hodnota veličiny ΔH pre prípad S1.

dokázala znížiť horný odhad energie základného stavu na rozdiel $E_{7P} - E_{\text{var}} = 0,008$ od presného výpočtu. Aj keď je to pod limitom Hartreeovej-Fockovej metódy, v tomto prípade je dostatočný priestor na zlepšenie. Na porovnanie ešte slúži experimentálna hodnota, ktorá sa líši od presného výpočtu (variačná metóda), pretože zahrňuje ďalšie korekcie ako relativistické korekcie, spinové interakcie, vplyv veľkosti jadra a mnoho ďalšie.

i	a	b	c	d	e	f	g	E	ΔE
0	0,1000	-2,000	-0,700	-107,0	-17,00	1,000	-7,000	-2,11881	0,00050
1	0,1019	-1,975	-0,687	-107,3	-15,53	0,826	-5,697	-2,13733	0,00050
1	0,0921	-2,102	-0,647	-107,4	-14,57	0,757	-5,511	-2,14497	0,00050
3	0,0878	-2,169	-0,512	-107,3	-12,48	0,803	-7,075	-2,15468	0,00050
4	0,0986	-2,092	-0,378	-107,1	-10,12	0,845	-8,201	-2,16429	0,00048
5	0,0995	-2,075	-0,405	-107,1	-10,48	0,844	-8,056	-2,16512	0,00049
6	0,0997	-2,073	-0,417	-107,1	-9,86	0,826	-8,076	-2,16550	0,00049
7	0,0984	-2,069	-0,429	-107,0	-7,24	0,760	-8,192	-2,16590	0,00049
8	0,0965	-2,083	-0,438	-107,0	-5,25	0,697	-8,125	-2,16617	0,00048
9	0,0971	-2,080	-0,435	-107,0	-5,86	0,712	-8,087	-2,16581	0,00049
10	0,0979	-2,077	-0,428	-107,0	-6,25	0,725	-8,086	-2,16639	0,00050
11	0,0987	-2,069	-0,422	-107,0	-5,63	0,712	-8,112	-2,16577	0,00049
12	0,0982	-2,073	-0,427	-107,0	-5,25	0,698	-8,090	-2,16577	0,00049
13	0,0977	-2,074	-0,431	-107,0	-4,63	0,679	-8,062	-2,16544	0,00049
14	0,0978	-2,073	-0,437	-107,0	-3,63	0,638	-8,090	-2,16571	0,00049
15	0,0978	-2,073	-0,436	-107,0	-3,02	0,608	-8,144	-2,16654	0,00049
16	0,0978	-2,073	-0,436	-107,0	-3,02	0,608	-8,144	-2,16719	0,00048

Tabuľka 3.3: Číslo kroku i , hodnoty parametrov po i -tom kroku optimalizácie, hodnoty energie E s odhadom chyby ΔE , viriálový pomer p s odhadom chyby Δp a hodnota veličiny ΔH pre prípad A1.

i	E	ΔE	p	Δp	ΔH
0	-2,11881	0,00050	-1,8808	0,0017	0,418
1	-2,13733	0,00050	-1,9987	0,0020	0,380
1	-2,14497	0,00050	-1,9289	0,0019	0,364
3	-2,15468	0,00050	-1,8793	0,0019	0,361
4	-2,16429	0,00048	-2,0004	0,0040	0,191
5	-2,16512	0,00049	-1,9961	0,0040	0,195
6	-2,16550	0,00049	-2,0019	0,0039	0,203
7	-2,16590	0,00049	-2,0022	0,0038	0,203
8	-2,16617	0,00048	-1,9825	0,0038	0,203
9	-2,16581	0,00049	-1,9922	0,0038	0,202
10	-2,16639	0,00050	-1,9891	0,0038	0,217
11	-2,16577	0,00049	-2,0138	0,0042	0,182
12	-2,16577	0,00049	-2,0102	0,0038	0,203
13	-2,16544	0,00049	-1,9986	0,0036	0,209
14	-2,16571	0,00049	-2,0027	0,0040	0,191
15	-2,16654	0,00049	-1,9999	0,0035	0,226
16	-2,16719	0,00048	-1,9971	0,0038	0,199

Tabuľka 3.4: Číslo kroku i , hodnoty energie E s odhadom chyby ΔE , viriálový pomer p s odhadom chyby Δp a hodnota veličiny ΔH pre prípad A1.

i	a	b	c	d	e	f	g	E	ΔE
0	0,1000	-0,700	-2,000	-107,00	-17,00	1,000	-7,000	-2,13525	0,00050
1	0,1071	-0,567	-1,996	-107,28	-15,86	0,860	-5,901	-2,16132	0,00050
2	0,1124	-0,488	-1,995	-107,04	-15,86	1,033	-7,489	-2,16564	0,00050
3	0,1138	-0,493	-1,993	-107,04	-15,48	1,018	-7,514	-2,16631	0,00049
4	0,1163	-0,496	-1,992	-107,03	-14,86	0,994	-7,568	-2,16592	0,00049
5	0,1235	-0,502	-1,987	-107,01	-13,25	0,937	-7,695	-2,16688	0,00049
6	0,1207	-0,497	-1,988	-107,01	-13,87	0,964	-7,681	-2,16698	0,00048
7	0,1224	-0,501	-1,987	-107,01	-13,48	0,945	-7,674	-2,16631	0,00048
8	0,1180	-0,496	-1,990	-107,00	-14,48	0,982	-7,616	-2,16700	0,00050

Tabuľka 3.5: Číslo kroku i , hodnoty parametrov po i -tom kroku optimalizácie, hodnoty energie E s odhadom chyby ΔE , viriálový pomer p s odhadom chyby Δp a hodnota veličiny ΔH pre prípad A2.

i	E	ΔE	p	Δp	ΔH
0	-2,13525	0,00050	-1,8645	0,0018	0,390
1	-2,16132	0,00050	-1,9971	0,0028	0,236
2	-2,16564	0,00050	-1,9834	0,0044	0,198
3	-2,16631	0,00049	-1,9869	0,0038	0,215
4	-2,16592	0,00049	-1,9992	0,0040	0,192
5	-2,16688	0,00049	-2,0002	0,0040	0,193
6	-2,16698	0,00048	-1,9892	0,0038	0,198
7	-2,16631	0,00048	-1,9955	0,0039	0,189
8	-2,16700	0,00050	-1,9863	0,0051	0,179

Tabuľka 3.6: Číslo kroku i , hodnoty energie E s odhadom chyby ΔE , viriálový pomer p s odhadom chyby Δp a hodnota veličiny ΔH pre prípad A2.

stav	1S He	3S He
model neinteragujúcich elektrónov	-2,75	—
blízko limity HFM ³ [6]	-2,8617	-2,1638
variálna metóda [7]	-2,903724	-2,175229
experiment [8], [9]	-2,903694	-2,175337
1 parameter	$(-2,8772 \pm 0,0004)$	—
7 parametrov	$(-2,9026 \pm 0,0005)$	$(-2,1672 \pm 0,0005)$

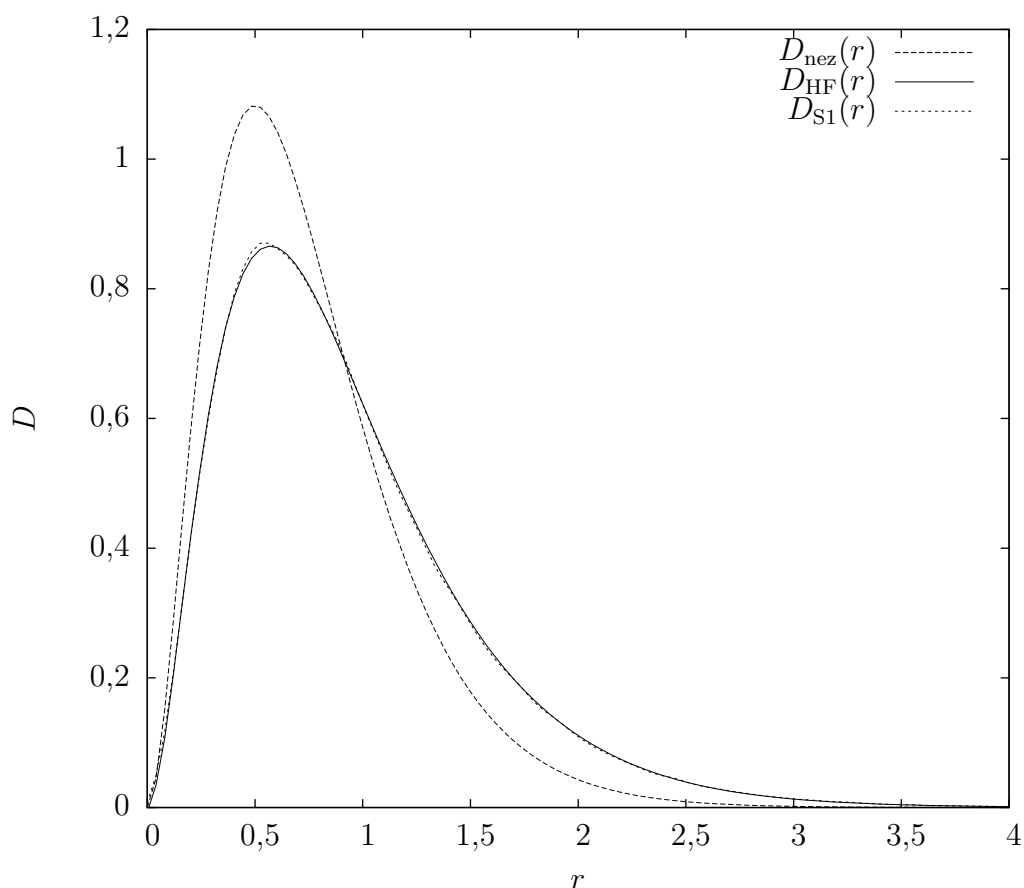
Tabuľka 3.7: Porovnanie vypočítaných a experimentálnych hodnôt.

3.1.5 Vlastnosti optimalizovanej funkcie

Okrem hodnoty energie, viriálneho pomeru a veličiny ΔH preskúmame ďalšie vlastnosti vlnovej funkcie s optimalizovanými parametrami a porovnáme ich s vlastnosťami skutočnej vlnovej funkcie.

Radiálna elektrónová hustota

Pre singlet sme v grafe 3.15 porovnali radiálnu elektrónovú hustotu neinteragujúcich elektrónov $D_{\text{nez}}(r)$, limity Hartreeovej-Fockovej metódy $D_{\text{HF}}(r)$ a našej vlnovej funkcie s optimalizovanými parametrami $D_{\text{S1}}(r)$. Na oboch hustotách $D_{\text{HF}}(r)$ a $D_{\text{S1}}(r)$ je vidno, že pridanie Coulombickej interakcie medzi elektróny spôsobilo rozostúpenie hustoty ďalej od centra. Obe hustoty $D_{\text{HF}}(r)$ a $D_{\text{S1}}(r)$ sa v grafe prekrývajú a nie je vidno žiadne výrazné rozdiely. I napriek tomu poskytuje optimalizovaná funkcia lepšiu hodnotu energie.



Obr. 3.15: Radiálna elektrónová hustota neinteragujúcich elektrónov $D_{\text{nez}}(r)$, limity Hartreeovej-Fockovej metódy $D_{\text{HF}}(r)$ a optimalizovanej vlnovej funkcie $D_{\text{S1}}(r)$ pre singletový stav.

Asymptotické správanie

Skutočná vlnová funkcia má podľa 1.2.7 spĺňať viaceré vlastnosti pri limitách $r_1 \rightarrow 0$, $r_{12} \rightarrow 0$ a $r_1 \rightarrow \infty$. Pri limitách do nuly pre náš systém musí platiť podľa

$$\begin{aligned}\lim_{r_1 \rightarrow 0} \frac{\frac{\partial \psi}{\partial r_1}}{\psi}(r_{mn}) &= -2, \\ \lim_{r_{12} \rightarrow 0} \frac{\frac{\partial \psi}{\partial r_{12}}}{\psi}(r_{mn}) &= \frac{1}{2}, \\ \lim_{r_1 \rightarrow \infty} \frac{\frac{\partial \psi}{\partial r_{12}}}{\psi}(r_{mn}) &= -\sqrt{2I}.\end{aligned}$$

Pre limitu do nekonečna sme použili v limite asymptotické správanie vlnovej funkcie a dostali uvedenú limitu.

Ionizačnú energiu určíme na základe rozdielu energie príslušného stavu z tabuľky 3.7 a energie He^+ , ktorá je v našom prípade presne 2,00. Tým dostávame pre singlet približne $\sqrt{2I_S} \approx 1,34$ a pre triplet približne $\sqrt{2I_A} \approx 0,59$.

Pre prípady S1 a A1 sme pre rôzne vlnové funkcie (rôzny krok optimalizácie i) vykreslili funkciu $\frac{\frac{\partial \psi}{\partial r_{jk}}}{\psi}(r_{mn})$ v závislosti od r_{jk} (ostatné závislé neurčené vzdialenosti boli zvolené jednotkové). Definujeme teda nasledujúce funkcie

$$\begin{aligned}F_i(r_1) &:= \frac{\frac{\partial \psi_i}{\partial r_1}}{\psi_i}(r_1, r_2 = 1, r_{12} = 1), \\ G_i(r_{12}) &:= \frac{\frac{\partial \psi_i}{\partial r_{12}}}{\psi_i}(r_1 = 1, r_2 = 1, r_{12}),\end{aligned}$$

kde vlnová funkcia $\psi_i(r_{mn})$ znamená vlnová funkcia po i -tom kroku optimalizácie.

V prípade limity do nekonečna sme vykreslili rovnakú funkciu v závislosti od $1/r_{jk}$.

V prípade singletu a limity jadro – elektrón (3.16) pozorujeme počas optimalizácie posun ku hodnote -2 . Pri limite elektrón – elektrón (3.17) dostaneme hodnotu 0,3 oproti 0,5. Pri limite do nekonečna (3.18) sa počas optimalizácie hodnota zvyšuje ku $-1,6$ (správna $-1,34$).

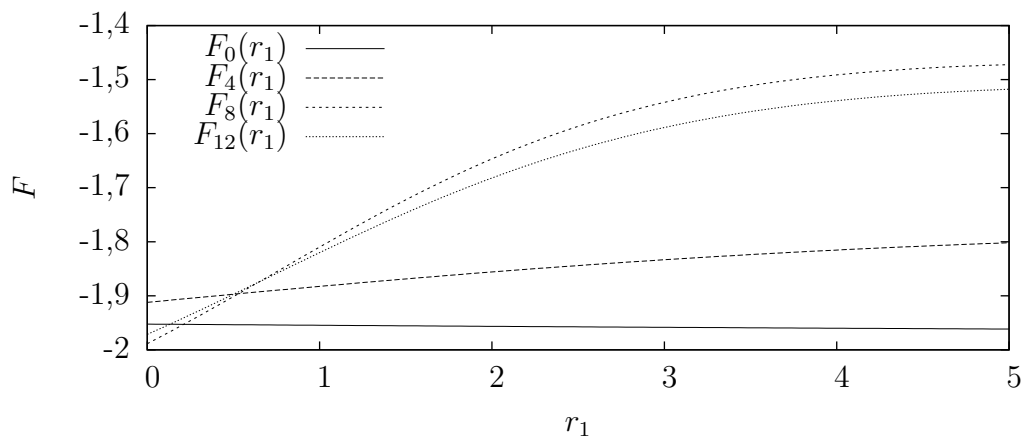
V prípade tripletu a limity jadro – elektrón (3.17) sa hodnota počas optimalizácie zastaví na $-1,7$. Limita elektrón – elektrón (3.18) dostávame podobne ako v prípade 3.17 hodnotu okolo 0,3. Pri limite do nekonečna (3.21) sa optimalizáciu dostávame ku hodnote $-0,45$ (správna $-0,59$).

V prípade limity do nekonečna nás príliš asymptotické správanie nezaujíma, keďže vplyv tejto oblasti do energie systému je malý. Preto táto limita nie je v prípade optimalizácie energie rozhodujúca a zhodná s teoretickou hodnotou.

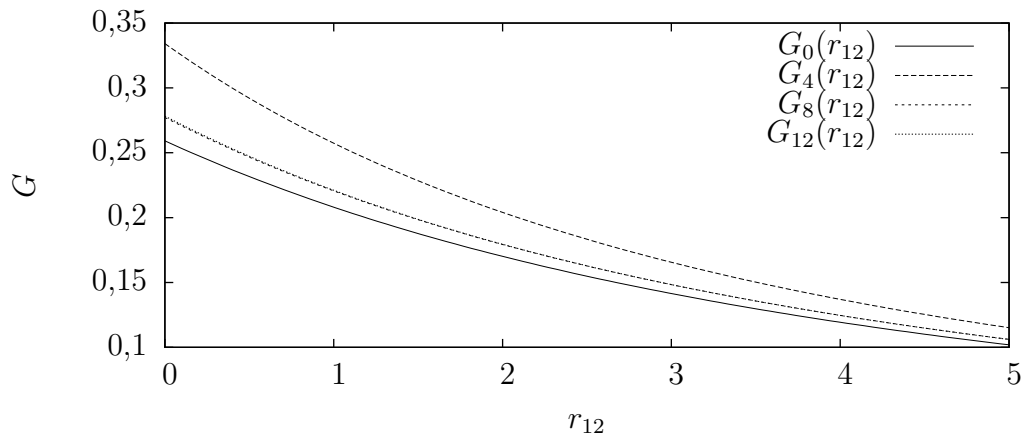
Všeobecne platí, že ak by sme parametre obmedzili podľa uvedených limitných podmienok, tak to zhorší výsledok energie, keďže hocijaká väzba znižuje stupeň voľnosti optimalizácie.

Prierez hustoty pravdepodobnosti

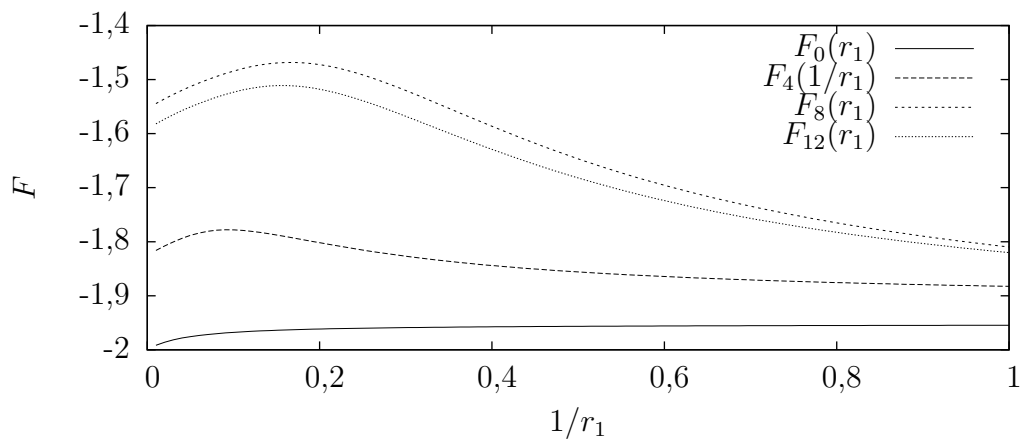
Pre lepšiu vizualizáciu sme si zobrazili v grafoch 3.22 až 3.27 rez kvadrátu vlnovej funkcie $|\psi|^2$ pre tri konkrétne polohy prvého elektrónu (poloha vo vzdialenosti 0,5; 1,0 a 1,5 od počiatku). Poloha 1. elektrónu je zobrazená plným krúžkom. Vykreslené sú kontúrové grafy kvadrátu vlnovej funkcie $|\psi|^2$ v rovine x_2, y_2 . V grafoch 3.22 až 3.24 vidíme presunutie hustoty pravdepodobnosti od



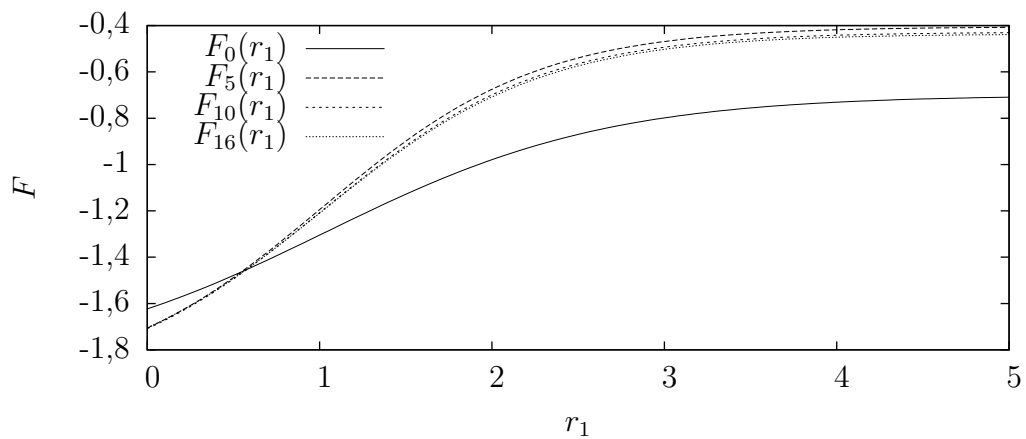
Obr. 3.16: Závislosť $F_i(r_1)$ pre 0., 4., 8. a 12. krok optimalizácie pre S1.



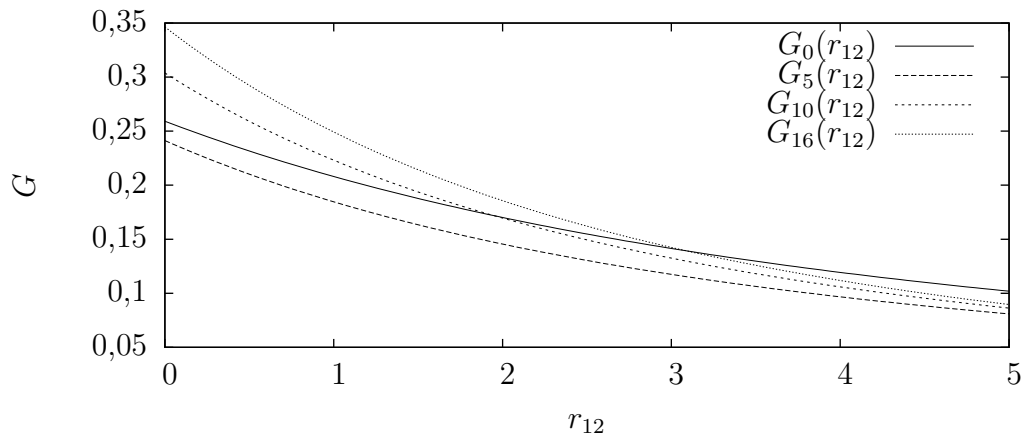
Obr. 3.17: Závislosť $G_i(r_{12})$ pre 0., 4., 8. a 12. krok optimalizácie pre S1.



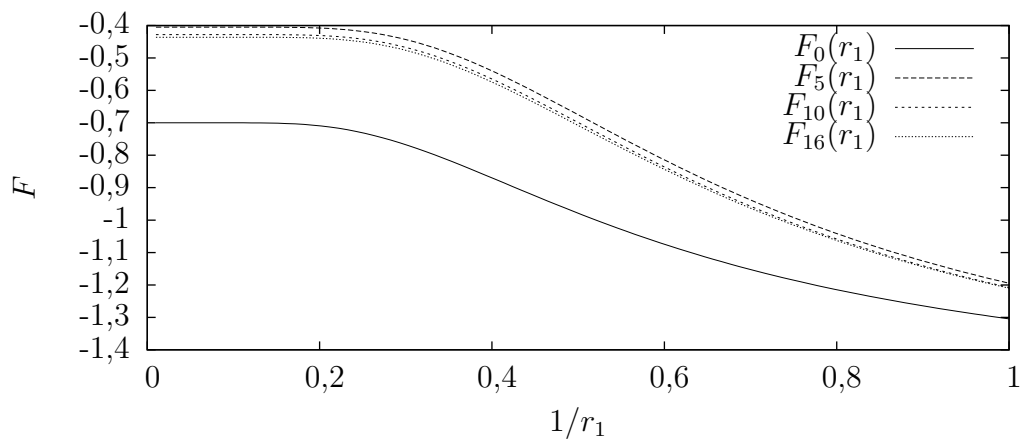
Obr. 3.18: Závislosť $F_i(r_1)$ pre 0., 4., 8. a 12. krok optimalizácie pre S1.



Obr. 3.19: Závislosť $F_i(r_1)$ pre 0., 5., 10. a 16. krok optimalizácie pre A1.

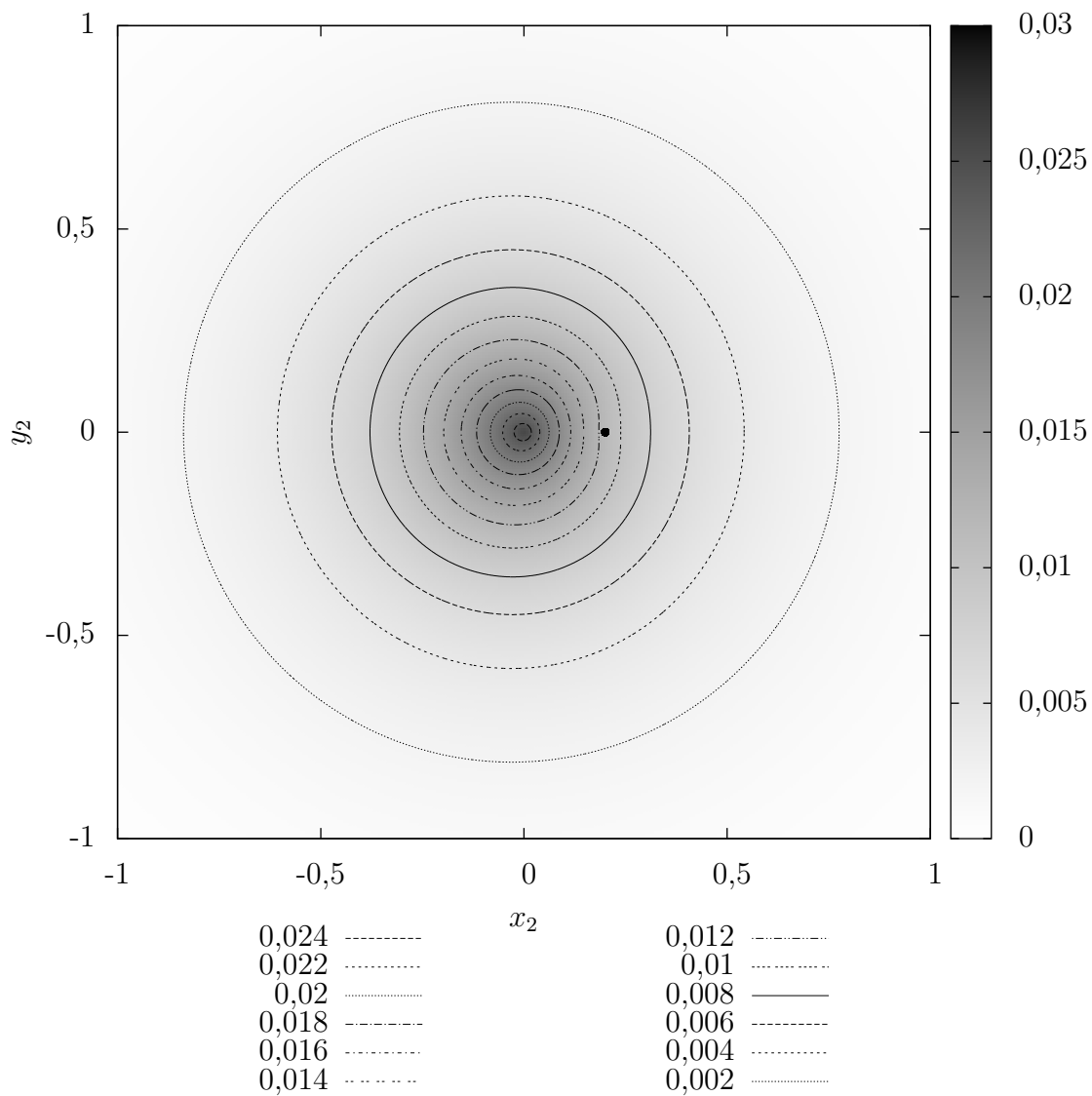


Obr. 3.20: Závislosť $G_i(r_{12})$ pre 0., 5., 10. a 16. krok optimalizácie pre A1.

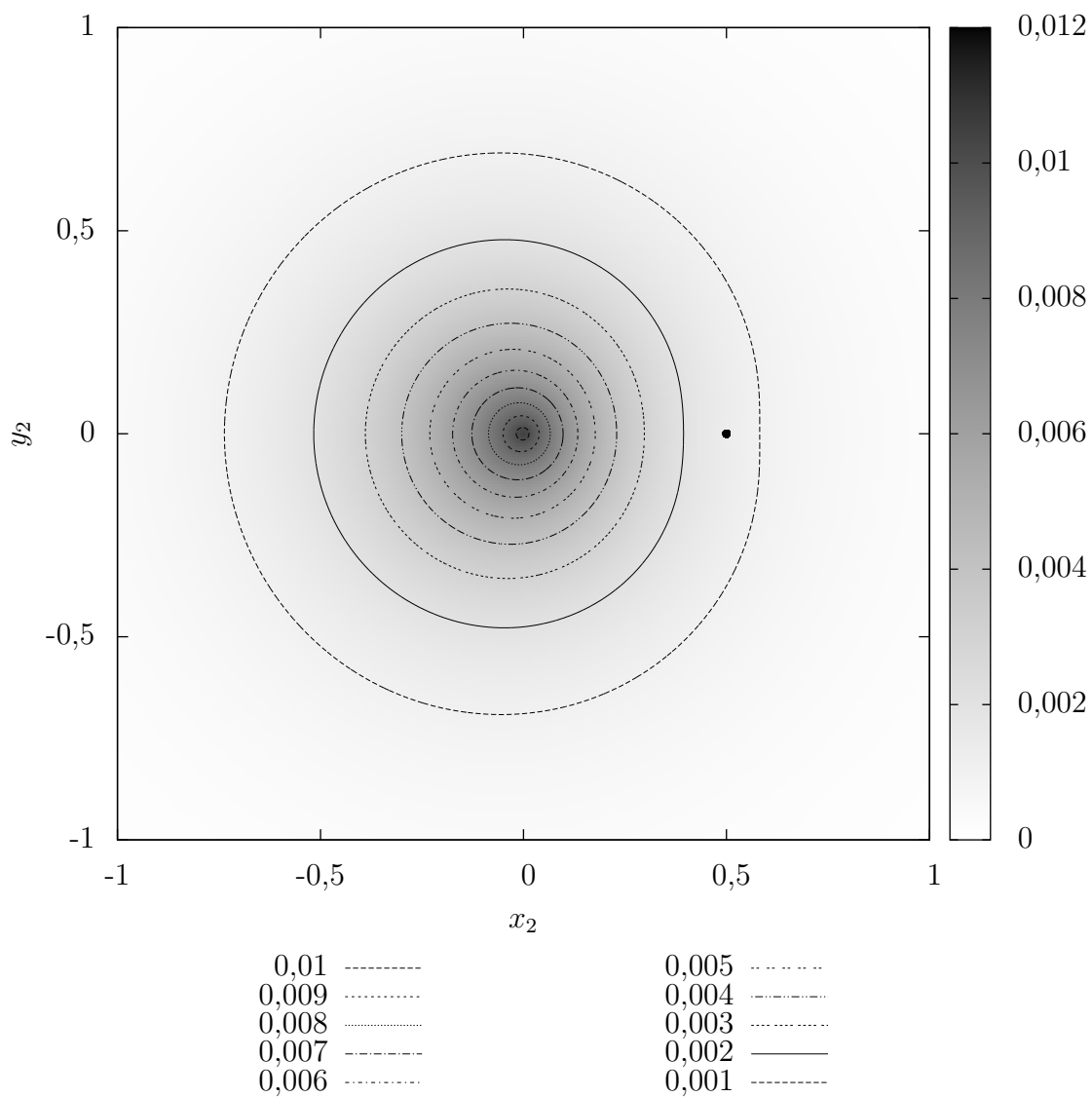


Obr. 3.21: Závislosť $F_i(r_1)$ pre 0., 5., 10. a 16. krok optimalizácie pre A1.

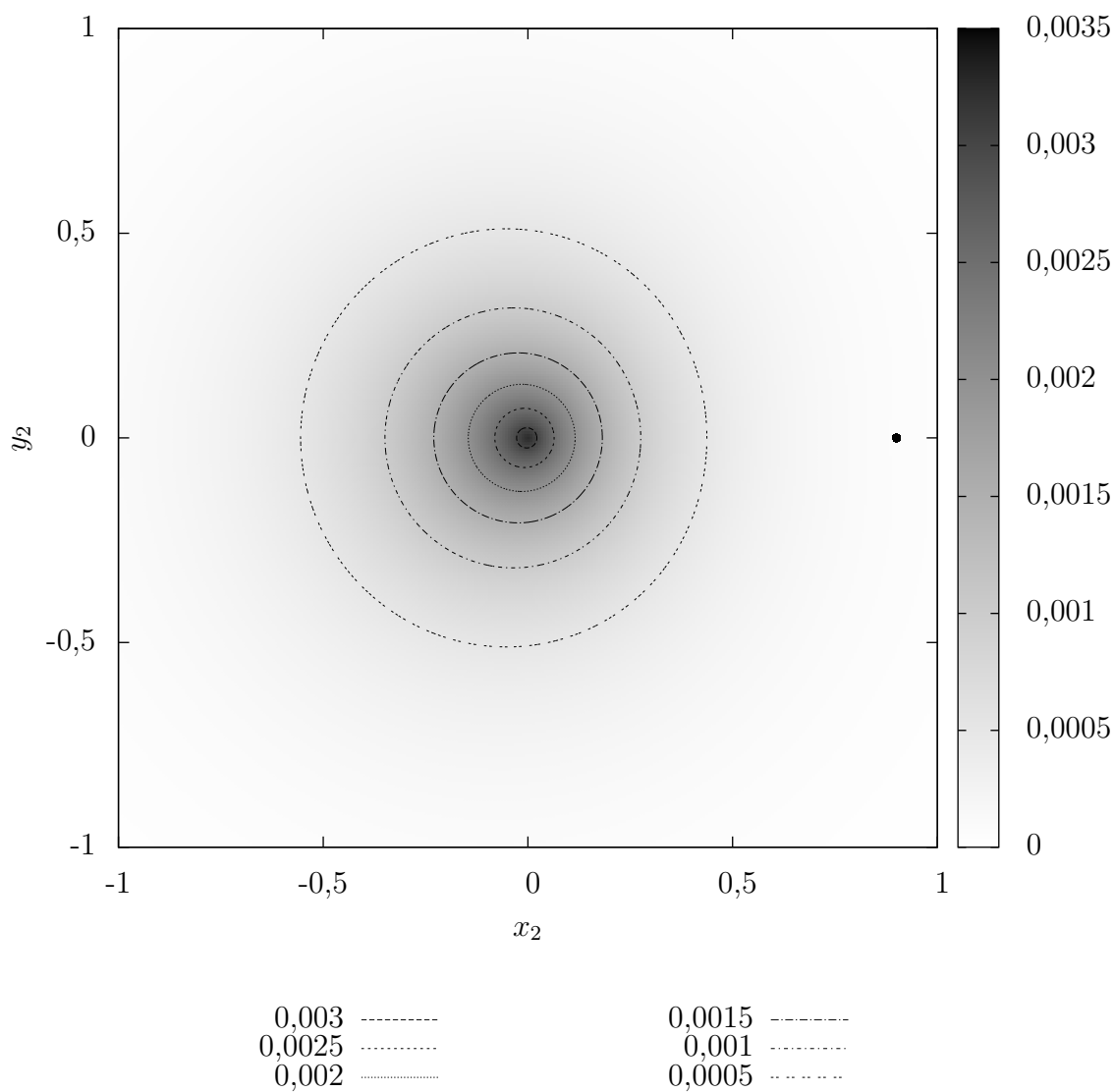
prvého elektrónu. V prípade tripletového stavu v grafoch 3.25 až 3.27 sa dokonca hustota pravdepodobnosti vytratila približne z oblasti polomeru vzdialenosti prvého elektrónu od počiatku.



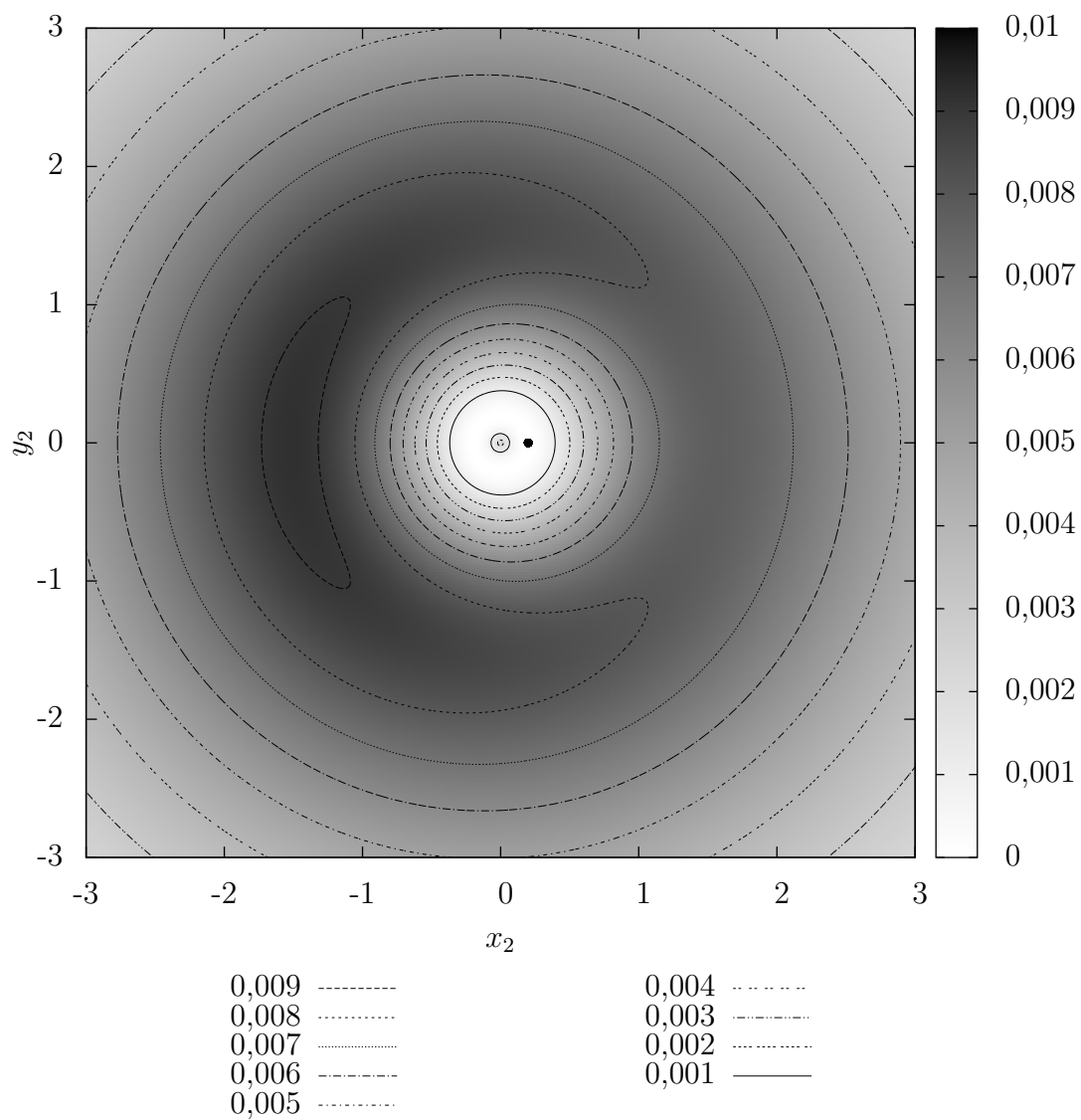
Obr. 3.22: Rez kvadrátu vlnovej funkcie $|\psi|^2$ optimalizovanej funkcie $|\psi_S(0, 2; 0; 0; x_2; y_2; 0)|^2$ pre singletový stav s vyzačením polohy 1. elektrónu (krúžok).



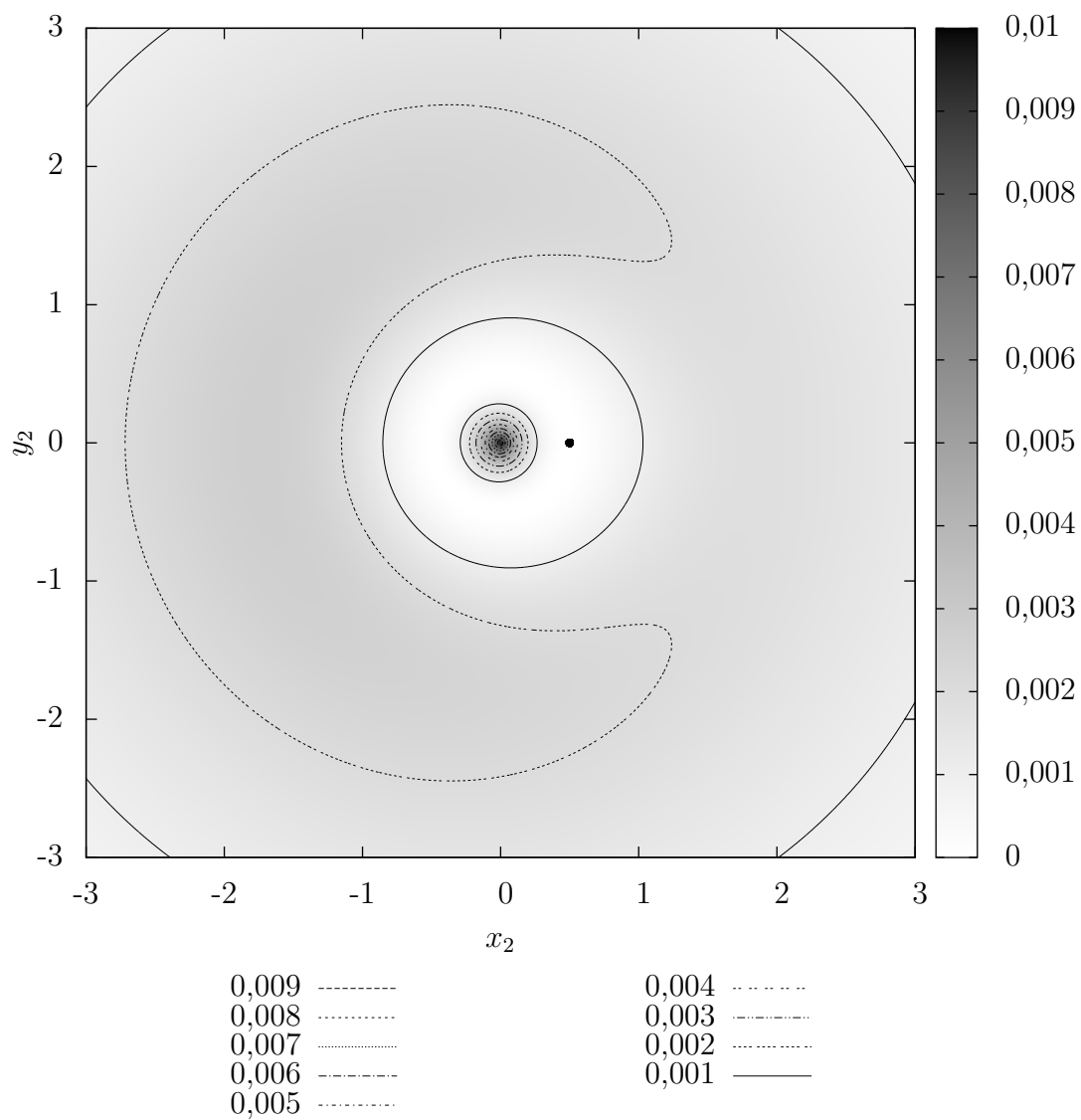
Obr. 3.23: Rez kvadrátu vlnovej funkcie $|\psi|^2$ optimalizovanej funkcie $|\psi_S(0, 5; 0; 0; x_2; y_2; 0)|^2$ pre singletový stav s vyzačením polohy 1. elektrónu (krúžok).



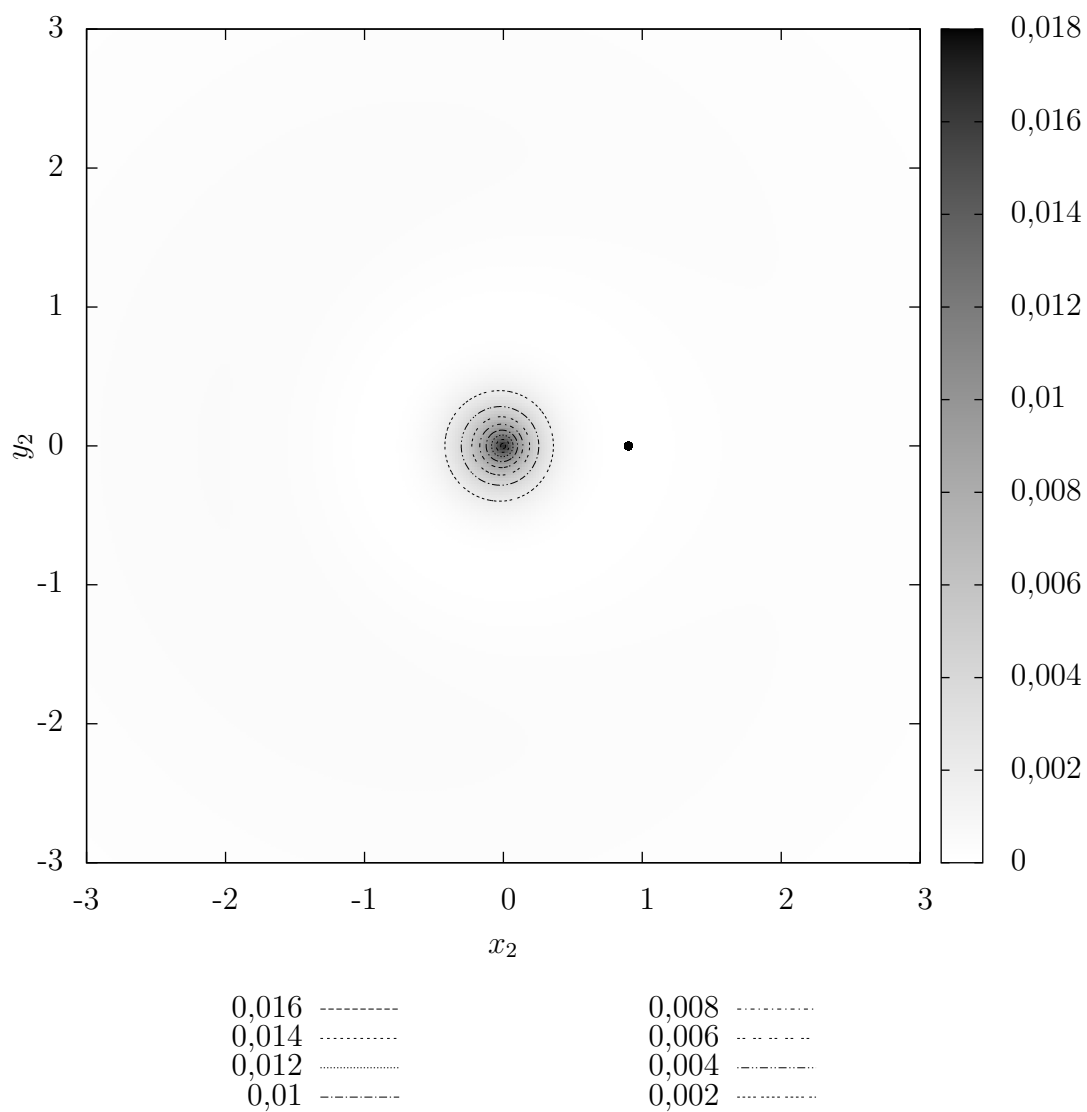
Obr. 3.24: Rez kvadrátu vlnovej funkcie $|\psi|^2$ optimalizovanej funkcie $|\psi_S(0, 9; 0; 0; x_2; y_2; 0)|^2$ pre singletový stav s vyzačením polohy 1. elektrónu (krúžok).



Obr. 3.25: Rez kvadrátu vlnovej funkcie $|\psi|^2$ optimalizovanej funkcie $|\psi_A(0, 2; 0; 0; x_2; y_2; 0)|^2$ pre tripletový stav s vyznačením polohy 1. elektrónu (krúžok).



Obr. 3.26: Rez kvadrátu vlnovej funkcie $|\psi|^2$ optimalizovanej funkcie $|\psi_A(0, 5; 0; 0; x_2; y_2; 0)|^2$ pre tripletový stav s vyzačením polohy 1. elektrónu (krúžok).



Obr. 3.27: Rez kvadrátu vlnovej funkcie $|\psi|^2$ optimalizovanej funkcie $|\psi_A(0, 9; 0; 0; x_2; y_2; 0)|^2$ pre tripletový stav s vyznačením polohy 1. elektrónu (krúžok).

Záver

Ukázalo sa, že pre jednoduchší systém, akým je hélium, sa dá nájsť skusmá vlnová funkcia, ktorá po optimalizácii pár parametrov dáva energiu s chemickou presnosťou. Keďže však naša metóda je variačná a jediná optimalizovaná vlastnosť je energia, tak optimalizovaná skusmá vlnová funkcia môže byť nezanedbateľne rozdielna od presnej vlnovej funkcie základného stavu. Tieto rozdiely je badať na niektorých vlastnostiach (asymptotické správanie, hodnota veličiny ΔH), ktoré sa nezhodujú s predpokladanými vlastnosťami presného základného stavu. Napriek tomu oproti Hartreeovej-Fockovej metóde sme dostali lepšie hodnoty odhadu energie. Na radiálnej hustote je však tento príspevok len relatívne málo viditeľný. Táto vlastnosť je rovnako badateľná i podľa [10].

4. Dodatky

4.1 Separácia premenných

Táto operácia sa často vyskytuje, preto ju urobíme všeobecne. Predpokladajme, že máme dva lineárne operátory \hat{A} a \hat{B} a riešime rovnicu vlastného problému s vlastným číslom λ vo tvare

$$(\hat{A} \pm \hat{B})f(X, Y) = \lambda f(X, Y), \quad (4.1)$$

kde X a Y sú usporiadané množiny premenných. Ďalej predpokladajme, že operátor \hat{A} pôsobí na funkciu $f(X, Y)$ iba podľa množiny premenných X a operátor \hat{B} pôsobí na funkciu $f(X, Y)$ iba podľa množiny premenných Y .

V takom prípade sa dá hľadaná funkcia $f(X, Y)$ riešiť v tvare

$$f(X, Y) = f_A(X)f_B(Y), \quad (4.2)$$

kde funkcie $f_A(X)$ a $f_B(Y)$ sú vlastné funkcie operátorov \hat{A} a \hat{B} s vlastnými číslami λ_A a λ_B

$$\hat{A}f_A(X) = \lambda_A f_A(X), \quad (4.3)$$

$$\hat{B}f_B(Y) = \lambda_B f_B(Y). \quad (4.4)$$

Overenie urobíme dosadením pričom využijeme linearitu operátorov, rovnice (4.3) a (4.4) a že pôsobia iba na svoju množinu premenných

$$\begin{aligned} (\hat{A} \pm \hat{B})f(X, Y) &= \hat{A}f(X, Y) \pm \hat{B}f(X, Y) = \\ &= \hat{A}(f_A(X)f_B(Y)) \pm \hat{B}(f_A(X)f_B(Y)) = \\ &= f_B(Y)\hat{A}f_A(X) \pm f_A(X)\hat{B}f_B(Y) = \\ &= f_B(Y)\lambda_A f_A(X) \pm f_A(X)\lambda_B f_B(Y) = \\ &= (\lambda_A \pm \lambda_B)f_A(X)f_B(Y) = \\ &= (\lambda_A \pm \lambda_B)f(X, Y). \end{aligned}$$

Vlastné číslo λ funkcie $f(X, Y)$ je potom súčtom/rozdielom vlastných čísel funkcií $f_A(X)$ a $f_B(Y)$

$$\lambda = \lambda_A \pm \lambda_B. \quad (4.5)$$

4.2 Ortogonalita vlastných funkcií

Predpokladajme operátor \hat{A} definovaný na množine funkcií M . Hermitovsky združený operátor k operátoru \hat{A} je operátor \hat{A}^+ tiež definovaný na množine funkcií M , pre ktorý platí

$$(\forall \psi \in M, \forall \varphi \in M) (\langle \psi | \hat{A}\varphi \rangle = \langle \hat{A}^+\psi | \varphi \rangle). \quad (4.6)$$

Ak platí $\hat{A} = \hat{A}^+$, tak operátor \hat{A} nazývame hermitovský (samozružený).

Jeho vlastné čísla sú reálne čísla, pričom sme v nasledujúcom postupe použili linearitu skalárneho súčinu $\langle \cdot, \cdot \rangle$ v druhom argumente, antilinearitu v prvom argumente, definíciu hermitovského operátora a rovnicu vlastného problému.

$$\begin{aligned}\langle \psi_n | \widehat{A}\psi_n \rangle &= \langle \psi_n | A_n\psi_n \rangle = \\ &= A_n \langle \psi_n | \psi_n \rangle \\ \langle \psi_n | \widehat{A}\psi_n \rangle &= \langle \widehat{A}^+\psi_n | \psi_n \rangle = \\ &= \langle \widehat{A}\psi_n | \psi_n \rangle = \\ &= \langle A_n\psi_n | \psi_n \rangle = \\ &= A_n^* \langle \psi_n | \psi_n \rangle \\ 0 &= (A_n - A_n^*) \langle \psi_n | \psi_n \rangle\end{aligned}$$

Norma funkcie ψ_n je nulová iba v prípade, ak je samotná funkcia nulová. Lenže pri riešení vlastného problému sa hľadajú netriviálne (nenulové) riešenia, preto musí platiť

$$A_n = A_n^*, \quad (4.7)$$

čo na obore komplexných čísel spĺňajú iba reálne čísla ($A_n \in \mathbb{R}$).

Predpokladajme, že máme dve vlastné funkcie ψ_1 a ψ_2 operátora \widehat{A} s vlastnými číslami A_1 a A_2 , ktoré sú rôzne ($A_1 \neq A_2$). Potom vlastné funkcie ψ_1 a ψ_2 sú ortogonálne voči skalárnemu súčinu $\langle \cdot, \cdot \rangle$.

$$\begin{aligned}\langle \psi_1 | \widehat{A}\psi_2 \rangle &= \langle \psi_1 | A_2\psi_2 \rangle = \\ &= A_2 \langle \psi_1 | \psi_2 \rangle \\ \langle \psi_1 | \widehat{A}\psi_2 \rangle &= \langle \widehat{A}\psi_1 | \psi_2 \rangle = \\ &= \langle A_1\psi_1 | \psi_2 \rangle = \\ &= A_1 \langle \psi_1 | \psi_2 \rangle \\ 0 &= (A_1 - A_2) \langle \psi_1 | \psi_2 \rangle\end{aligned}$$

Keďže sme predpokladali rôzne vlastné čísla, musí byť nulový skalárny súčin vlnových funkcií ψ_1 a ψ_2 .

4.3 Komutátory

Pri odvádzaní viriálového teorému bolo potrebné vypočítať dva komutátory. Preto si v tomto dodatku najprv napíšeme základné vlastnosti komutátorov (ktoré použijeme) a dané komutátory vyjadríme.

Vo všeobecnosti definujeme *komutátor* $[\cdot, \cdot]$ ako bilinéarne zobrazenie na vektorovom priestore W do vektorového priestoru W

$$[\cdot, \cdot] : W \times W \rightarrow W, \quad (4.8)$$

ktoré spĺňa dve vlastnosti:

1. Pre každé dva prvky a a b z W je *antisymetrický*:

$$(\forall a, b \in W) ([a, b] = -[b, a]). \quad (4.9)$$

2. Pre každé tri prvky a, b a c z W platí *Jacobiho identita*:

$$(\forall a, b, c \in W) ([a, b], c] + [[b, c], a] + [[c, a], b] = 0) . \quad (4.10)$$

Dvojica $(W, [\cdot, \cdot])$ sa potom označuje ako *Lieova algebra*.

Voľba komutátora v tvare

$$[a, b] := ab - ba \quad (4.11)$$

spĺňa definíciu a je najčastejšia.

Najskôr si zopár jednoduchých komutátorov vypočítame, neskôr ich využijeme. Samotné komutátory sú operátory a sú teda definované pôsobením na nejakú všeobecnú funkciu v danej reprezentácii, napríklad súradnicovej $f(\mathbf{r}_j)$. Základné operátory, v ktorých budeme vyjadrovať operátory sú operátor k -tej súradnice \hat{x}_k , k -tej združenej hybnosti \hat{p}_k a potenciálu \hat{V} . Ich vzájomné komutátory vyrátame. Komutátor rovnakého operátora je z antisymetrie nulový operátor (nula, vyčíslený na funkcii)

$$[\hat{A}, \hat{A}] = -[\hat{A}, \hat{A}] = \hat{0} = 0 . \quad (4.12)$$

Vďaka antisymetrii tak ľahko vypočítame aj komutátory, ktoré vzniknú zámennou operátorov. Preto budeme rátať iba päť netriviálnych komutátorov: $[\hat{x}_k, \hat{x}_l]$, $[\hat{p}_k, \hat{p}_l]$, $[\hat{x}_k, \hat{p}_l]$, $[\hat{x}_k, \hat{V}]$ a $[\hat{p}_k, \hat{V}]$.

V súradnicovej reprezentácii majú operátory tvar

$$\begin{aligned} \hat{x}_k &= x_k , \\ \hat{p}_k &= -i\hbar \frac{\partial}{\partial x_k} , \\ \hat{V} &= V(\mathbf{r}_j) , \end{aligned}$$

kde potenciál $V(\mathbf{r}_j)$ explicitne závisí iba od súradníc častíc (\mathbf{r}_j) .

V impulzovej reprezentácii majú operátory tvar

$$\begin{aligned} \hat{p}_k &= p_k , \\ \hat{x}_k &= i\hbar \frac{\partial}{\partial p_k} , \\ \hat{V} &= V \left(i\hbar \frac{\partial}{\partial p_j} \right) . \end{aligned}$$

Komutátory $[\hat{x}_k, \hat{x}_l]$ a $[\hat{p}_k, \hat{p}_l]$ sú nakoniec pomerne jednoduché pri vhodne zvolenej reprezentácii. V prvom prípade je vhodná súradnicová reprezentácia, kde je komutátor nulový jednoducho z komutácie násobenia funkcií, v druhom prípade impulzová reprezentácia, kde je rovnako komutátor nulový z rovnakého dôvodu. V opačne zvolenej reprezentácii (v prvom prípade impulzovej, v druhom prípade súradnicovej) dostávame rovnako nulu za predpokladu zameniteľnosti (komutácie) parciálnych derivácií, čo platí pre funkcie triedy $\mathcal{C}^2(\mathbb{R}^{3N})$ ¹.

¹Priestor funkcií $\mathcal{C}^n(M)$ obsahuje funkcie, ktoré majú na množine M spojité všetky n -té parciálne derivácie, na hranici množiny M v príslušnom smere.

Pre tretí komutátor dostávame teda

$$\begin{aligned} [\hat{x}_k, \hat{p}_l]f &= x_k \left(-i\hbar \frac{\partial f}{\partial x_l} \right) - \left[-i\hbar \frac{\partial}{\partial x_l} (x_k f) \right] = \\ &= -i\hbar x_k \frac{\partial f}{\partial x_l} + i\hbar \delta_{kl} f + i\hbar x_k \frac{\partial f}{\partial x_l} = \\ &= i\hbar \delta_{kl} f, \end{aligned}$$

kde sme použili Leibnizovo pravidlo pri derivácii a fakt, že súradnice sú nezávislé.²
Pre štvrtý komutátor máme

$$[\hat{x}_k, \hat{V}]f = x_k (Vf) - V(x_k f) = 0,$$

kde sme dostali nulu, lebo operácie sú len násobenia funkcií, ktoré je komutatívne.
A posledný komutátor sa rovná

$$\begin{aligned} [\hat{p}_k, \hat{V}]f &= -i\hbar \frac{\partial}{\partial x_k} (Vf) - V \left(-i\hbar \frac{\partial f}{\partial x_k} \right) = \\ &= -i\hbar \frac{\partial V}{\partial x_k} f - i\hbar V \frac{\partial f}{\partial x_k} + i\hbar V \frac{\partial f}{\partial x_k} = \\ &= -i\hbar \frac{\partial V}{\partial x_k} f, \end{aligned}$$

kde sme použili Leibnizovo pravidlo pri derivácii.
Dostali sme teda päť užitočných komutátorov

$$[\hat{x}_k, \hat{x}_l] = 0, \quad (4.13)$$

$$[\hat{p}_k, \hat{p}_l] = 0, \quad (4.14)$$

$$[\hat{x}_k, \hat{p}_l] = i\hbar \delta_{kl}, \quad (4.15)$$

$$[\hat{x}_k, \hat{V}] = 0, \quad (4.16)$$

$$[\hat{p}_k, \hat{V}] = -i\hbar \frac{\partial V}{\partial x_k}. \quad (4.17)$$

Pri výpočte komutátora $[\hat{x}_k, \hat{H}]$ použijeme nerelativistický Hamiltonián, ktorý má tvar

$$\hat{H} = \sum_{l=1}^{3N} \frac{\hat{p}_l^2}{2m_l} + \hat{V}. \quad (4.18)$$

Dosadíme rovnicu (4.18) do komutátora, využijeme linearitu komutátora z

²Pri parciálnej derivácii súradnice x_k podľa súradnice x_l je v prípade zhody indexov jednotka a pri rôznych indexoch nula, kvôli nezávislosti súradníc.

definície, vzťah (1.59) a komutátory (4.15) a (4.16).

$$\begin{aligned}
[\hat{x}_k, \hat{H}] &= \left[\hat{x}_k, \sum_{l=1}^{3N} \frac{\hat{p}_l^2}{2m_l} + \hat{V} \right] = \\
&= \sum_{l=1}^{3N} \frac{1}{2m_l} [\hat{x}_k, \hat{p}_l^2] + [\hat{x}_k, \hat{V}] = \\
&= \sum_{l=1}^{3N} \frac{1}{2m_l} (\hat{p}_l [\hat{x}_k, \hat{p}_l] + [\hat{x}_k, \hat{p}_l] \hat{p}_l) + 0 = \\
&= \sum_{l=1}^{3N} \frac{1}{2m_l} (2i\hbar\delta_{kl}\hat{p}_l) = \\
&= \frac{i\hbar}{m_l} \hat{p}_k.
\end{aligned}$$

Podobne postupujeme aj v prípade výpočtu komutátora $[\hat{p}_k, \hat{H}]$. Dosadíme rovnicu (4.18) do komutátora, využijeme linearitu komutátora z definície, vzťah (1.59) a komutátory (4.14) a (4.17).

$$\begin{aligned}
[\hat{p}_k, \hat{H}] &= \left[\hat{p}_k, \sum_{l=1}^{3N} \frac{\hat{p}_l^2}{2m_l} + \hat{V} \right] = \\
&= \sum_{l=1}^{3N} \frac{1}{2m_l} [\hat{p}_k, \hat{p}_l^2] + [\hat{p}_k, \hat{V}] = \\
&= \sum_{l=1}^{3N} \frac{1}{2m_l} (\hat{p}_l [\hat{p}_k, \hat{p}_l] + [\hat{p}_k, \hat{p}_l] \hat{p}_l) + [\hat{p}_k, \hat{V}] = \\
&= 0 - i\hbar \frac{\partial V}{\partial x_k}.
\end{aligned}$$

Dostali sme tak dva užitočné komutátory

$$[\hat{x}_k, \hat{H}] = \frac{i\hbar}{m_l} \hat{p}_k, \quad (4.19)$$

$$[\hat{p}_k, \hat{H}] = -i\hbar \frac{\partial V}{\partial x_k}. \quad (4.20)$$

4.4 Potenciál vo viacčasticovom systéme

V prvom rade budeme musieť zaviesť okrem indexovania súradníc (k, l, \dots) i indexovanie častíc (K, L, \dots) . Kartézske súradnice K -tej častice sú

$$(X_K, Y_K, Z_K) = (x_{3K-2}, x_{3K-1}, x_{3K}). \quad (4.21)$$

Ak je potenciál medzi dvoma časticami (K -ta a L -tá častica) vzájomne sféricky symetrický, závisí iba od ich vzájomnej vzdialenosti r_{KL} danej vzťahom (*euklidovská metrika*)

$$r_{KL} = \left[(X_K - X_L)^2 + (Y_K - Y_L)^2 + (Z_K - Z_L)^2 \right]^{1/2}, \quad (4.22)$$

a potenciál medzi dvoma časticami má potom tvar

$$V_{KL}(X_K, Y_K, Z_K, X_L, Y_L, Z_L) = V_{KL}(r_{KL}). \quad (4.23)$$

Nech navyše má ešte potenciál špeciálny mocninový tvar

$$V_{KL}(r_{KL}) = \alpha_{KL} r_{KL}^n, \quad (4.24)$$

kde n je exponent pre všetky dvojice (K, L) rovnaký a α_{KL} koeficient vo všeobecnosti závisujúci od dvojice (K, L) .

Celkový potenciál $V(\mathbf{r}_j)$ potom dostaneme sčítaním, cez všetky dvojice

$$V(\mathbf{r}_j) = \sum_{K=1}^N \sum_{L=K+1}^N V_{KL}(r_{KL}). \quad (4.25)$$

Nás bude zaujímať derivácia potenciálu podľa súradnice x_k , teda $\frac{\partial V}{\partial x_k}$, ale potenciál máme vyjadrený v súradniciach (X_K, Y_K, Z_K) . Preto budeme musieť poznať vzťah medzi nimi pri derivovaní. Celkom jednoducho dôjdeme na vzťahy

$$\frac{\partial X_K}{\partial x_k} = \delta_{k(3K-2)}, \quad \frac{\partial Y_K}{\partial x_k} = \delta_{k(3K-1)}, \quad \frac{\partial Z_K}{\partial x_k} = \delta_{k(3K)}, \quad (4.26)$$

kde sme využili to, že ak ide o korešpondujúce súradnice (rovnaké) derivácia je jeden a v inom prípade sú to nezávislé súradnice, vtedy je derivácia nula.

Pri rátaní $\frac{\partial V}{\partial x_k}$ tak dostaneme

$$\begin{aligned} \frac{\partial V}{\partial x_k} &= \frac{\partial}{\partial x_k} \left(\sum_{K=1}^N \sum_{L=K+1}^N V_{KL}(r_{KL}) \right) = \\ &= \sum_{K=1}^N \sum_{L=K+1}^N \frac{\partial}{\partial x_k} (V_{KL}(r_{KL})) = \\ &= \sum_{K=1}^N \sum_{L=K+1}^N \frac{\partial}{\partial x_k} (\alpha_{KL} r_{KL}^n) = \\ &= \sum_{K=1}^N \sum_{L=K+1}^N \alpha_{KL} \frac{\partial}{\partial x_k} \left[(X_K - X_L)^2 + (Y_K - Y_L)^2 + (Z_K - Z_L)^2 \right]^{n/2} = \\ &= \sum_{K=1}^N \sum_{L=K+1}^N \alpha_{KL} \frac{n}{2} \left[(X_K - X_L)^2 + (Y_K - Y_L)^2 + (Z_K - Z_L)^2 \right]^{n/2-1} \cdot \\ &\quad \cdot 2 \left[(X_K - X_L)(\delta_{k(3K-2)} - \delta_{k(3L-2)}) + (Y_K - Y_L)(\delta_{k(3K-1)} - \delta_{k(3L-1)}) + \right. \\ &\quad \left. + (Z_K - Z_L)(\delta_{k(3K)} - \delta_{k(3L)}) \right]. \end{aligned}$$

Nás zaujíma výraz $\nabla V \cdot \hat{r}$. Teda do definície v súradnicovej reprezentácii do-

sadíme predchádzajúci člen a už len upravujeme

$$\begin{aligned}
\nabla V \cdot \hat{\mathbf{r}} &= \sum_{k=1}^{3N} \frac{\partial V}{\partial x_k} x_k = \\
&= n \sum_{k=1}^{3N} \sum_{K=1}^N \sum_{L=K+1}^N \alpha_{KL} \left[(X_K - X_L)^2 + (Y_K - Y_L)^2 + (Z_K - Z_L)^2 \right]^{n/2-1} \cdot \\
&\quad \cdot \left[(X_K - X_L)(\delta_{k(3K-2)} - \delta_{k(3L-2)}) + (Y_K - Y_L)(\delta_{k(3K-1)} - \delta_{k(3L-1)}) + \right. \\
&\quad \left. + (Z_K - Z_L)(\delta_{k(3K)} - \delta_{k(3L)}) \right] x_k = \\
&= n \sum_{K=1}^N \sum_{L=K+1}^N \alpha_{KL} \left[(X_K - X_L)^2 + (Y_K - Y_L)^2 + (Z_K - Z_L)^2 \right]^{n/2-1} \cdot \\
&\quad \cdot \sum_{k=1}^{3N} \left[(X_K - X_L)(\delta_{k(3K-2)} - \delta_{k(3L-2)}) + (Y_K - Y_L)(\delta_{k(3K-1)} - \delta_{k(3L-1)}) + \right. \\
&\quad \left. + (Z_K - Z_L)(\delta_{k(3K)} - \delta_{k(3L)}) \right] x_k = \\
&= n \sum_{K=1}^N \sum_{L=K+1}^N \alpha_{KL} \left[(X_K - X_L)^2 + (Y_K - Y_L)^2 + (Z_K - Z_L)^2 \right]^{n/2-1} \cdot \\
&\quad \cdot \left[(X_K - X_L)(x_{3K-2} - x_{3L-2}) + (Y_K - Y_L)(x_{3K-1} - x_{3L-1}) + \right. \\
&\quad \left. + (Z_K - Z_L)(x_{3K} - x_{3L}) \right] = \\
&= n \sum_{K=1}^N \sum_{L=K+1}^N \alpha_{KL} \left[(X_K - X_L)^2 + (Y_K - Y_L)^2 + (Z_K - Z_L)^2 \right]^{n/2} = \\
&= n \hat{V},
\end{aligned}$$

kde sme využili zameniteľnosť súm, zosumovanie cez Kroneckerovo delta, vzájomné jednoznačné priradenie medzi súradnicami $X_K, X_L, Y_K, Y_L, Z_K, Z_L$ a $x_{3K-2}, x_{3L-2}, x_{3K-1}, x_{3L-1}, x_{3K}, x_{3L}$, čím sme dostali znova výraz pôvodný potenciál \hat{V} .

Dostali sme teda, že v tomto špeciálnom prípade vzájomne sféricky symetrického potenciálu medzi časticami s daným exponentom n má operátor $\nabla V \cdot \hat{\mathbf{r}}$ tvar

$$\nabla V \cdot \hat{\mathbf{r}} = n \hat{V}. \quad (4.27)$$

4.5 Štatistika

Pri odvádzaní štatistických vlastností estimátorov si najprv zadefinujeme základné pojmy. *Stredná hodnota* veličiny Y na hustote pravdepodobnosti $\varrho(\mathbf{r})$ v priestore V je (v prípade spojitého rozdelenia) definovaná ako

$$\mathcal{E}(Y) := \int_V \varrho(\mathbf{r}) Y(\mathbf{r}) dV. \quad (4.28)$$

Rozptyl (variancia) veličiny Y na hustote pravdepodobnosti $\varrho(\mathbf{r})$ v priestore V je definovaný ako

$$\mathcal{V}ar(Y) := \mathcal{E} \left((Y - \mathcal{E}(Y))^2 \right). \quad (4.29)$$

Vďaka definícii strednej hodnoty je v argumente stredná hodnota lineárna, lebo aj integrál je lineárny

$$\begin{aligned}\mathcal{E}(Y + \alpha Z) &= \int_V \varrho(\mathbf{r}) (Y + \alpha Z)(\mathbf{r}) dV = \\ &= \int_V \varrho(\mathbf{r}) Y(\mathbf{r}) dV + \alpha \int_V \varrho(\mathbf{r}) Z(\mathbf{r}) dV = \\ &= \mathcal{E}(Y) + \alpha \mathcal{E}(Z),\end{aligned}$$

kde $\alpha \in \mathbb{C}$ je konštanta.

Pre linearitu strednej hodnoty potom dostávame pre rozptyl z definície vzťah

$$\mathcal{V}ar(Y) = \mathcal{E}\left((Y - \mathcal{E}(Y))^2\right) = \mathcal{E}\left(Y^2 - 2\mathcal{E}(Y)Y + \mathcal{E}(Y)^2\right) = \mathcal{E}(Y^2) - \mathcal{E}(Y)^2.$$

Ak máme dve veličiny Y a Z , každá generovaná na vlastnom priestore V a W , tak vo všeobecnosti máme hustotu pravdepodobnosti $\tilde{\varrho}(\mathbf{r}, \mathbf{s})$ závislú na oboch hodnotách $\mathbf{r} \in V$ a $\mathbf{s} \in W$. Ak sú však tieto veličiny nezávislé (Y a Z , resp. \mathbf{r} a \mathbf{s}), potom existujú hustoty pravdepodobnosti $\varrho_1(\mathbf{r})$ a $\varrho_2(\mathbf{s})$, ktoré sú normalizované na svojich priestoroch (V , resp. W)

$$\int_V \varrho_1(\mathbf{r}) dV = 1 \quad \int_W \varrho_2(\mathbf{s}) dW = 1$$

a platí (vo význame rovnosti funkcií)

$$\tilde{\varrho}(\mathbf{r}, \mathbf{s}) = \varrho_1(\mathbf{r})\varrho_2(\mathbf{s}). \quad (4.30)$$

Potom má stredná hodnota súčinu veličín Y a Z nasledujúcu vlastnosť

$$\begin{aligned}\mathcal{E}(YZ) &= \int_{V \times W} \tilde{\varrho}(\mathbf{r}, \mathbf{s}) (YZ)(\mathbf{r}, \mathbf{s}) dV dW = \\ &= \int_V \varrho_1(\mathbf{r}) Y(\mathbf{r}) \left(\int_W \varrho_2(\mathbf{s}) Z(\mathbf{s}) dW \right) dV = \mathcal{E}(Y) \mathcal{E}(Z).\end{aligned} \quad (4.31)$$

Pri našom výpočte máme zopár predpokladov. Predpokladáme, že máme dané rozdelenie $\varrho(\mathbf{r})$, na ktorom máme veličinu $A(\mathbf{r})$, ktorá má strednú hodnotu μ_A a rozptyl σ_A^2 . Potom predpokladáme, že vzorky \mathbf{r}_i sú z rozdelenia $\varrho(\mathbf{r})$, a teda hodnoty veličiny A na vzorkách $A_i = A(\mathbf{r}_i)$ sú z rozdelenia so strednou hodnotou μ_A a rozptylom σ_A^2

$$\mathcal{E}(A_i) = \mathcal{E}(A) = \mu_A, \quad (4.32)$$

$$\mathcal{V}ar(A_i) = \mathcal{V}ar(A) = \sigma_A^2. \quad (4.33)$$

Ďalej predpokladáme, že vzorky \mathbf{r}_i sú navzájom nezávislé, preto podľa (4.31) platí

$$\mathcal{E}(A_i A_j) = \mathcal{E}(A_i) \mathcal{E}(A_j) = \mu_A^2, \quad (4.34)$$

pri voľbe rôznych veličín (teda $i \neq j$).

Hodnotu pri voľbe rovnakých veličín dostaneme z definície rozptylu

$$\begin{aligned}\mathcal{V}ar(A_i) &= \mathcal{E}(A_i^2) - \mathcal{E}(A_i)^2 = \mathcal{E}(A_i^2) - \mu_A^2 = \sigma_A^2, \\ \mathcal{E}(A_i^2) &= \sigma_A^2 + \mu_A^2.\end{aligned} \quad (4.35)$$

Vypočítame strednú hodnotu a rozptyl výberového priemeru, pričom využijeme jeho definíciu (2.7), linearitu strednej hodnoty a predpoklad (4.32)

$$\begin{aligned}\mathcal{E}(\tilde{A}) &= \mathcal{E}\left(\frac{1}{N}\sum_{j=1}^N A_j\right) = \frac{1}{N}\sum_{j=1}^N \mathcal{E}(A_j) = \\ &= \frac{1}{N}\sum_{j=1}^N \mu_A = \frac{N}{N}\mu_A = \mu_A.\end{aligned}$$

Pri rozptyle postupujeme obdobne

$$\begin{aligned}\mathcal{V}ar(\tilde{A}) &= \mathcal{E}\left(\left(\tilde{A} - \mathcal{E}(\tilde{A})\right)^2\right) = \mathcal{E}\left(\left(\frac{1}{N}\sum_{j=1}^N A_j - \mu_A\right)^2\right) = \\ &= \mathcal{E}\left(\left(\frac{1}{N}\sum_{j=1}^N A_j\right)^2 - \frac{2}{N}\mu_A\sum_{j=1}^N A_j + \mu_A^2\right) = \\ &= \frac{1}{N^2}\sum_{j=1}^N\sum_{k=1}^N \mathcal{E}(A_j A_k) - \frac{2}{N}\mu_A\sum_{j=1}^N \mathcal{E}(A_j) + \mu_A^2 = \\ &= \frac{1}{N^2}\left(N(\sigma_A^2 + \mu_A^2) + (N^2 - N)\mu_A^2\right) - \frac{2}{N}N\mu_A^2 + \mu_A^2 = \\ &= \frac{\sigma_A^2}{N},\end{aligned}$$

kde sme využili linearitu strednej hodnoty a výrazy (4.32), (4.34) a (4.35).

Ďalej potrebujeme zistiť, či sú naše estimátory vychýlené (či v priemere dávajú odhad totožný s odhadovanou veličinou). Najprv vypočítame strednú hodnotu estimátora $\tilde{\sigma}_A^2$

$$\begin{aligned}\mathcal{E}(\tilde{\sigma}_A^2) &= \mathcal{E}\left(\frac{1}{N}\sum_{j=1}^N (A_j - \mu_A)^2\right) = \\ &= \frac{1}{N}\mathcal{E}\left(\sum_{j=1}^N A_j^2 - 2\mu_A\sum_{j=1}^N A_j + \sum_{j=1}^N \mu_A^2\right) = \\ &= \frac{1}{N}\left(\sum_{j=1}^N \mathcal{E}(A_j^2) - 2\mu_A\sum_{j=1}^N \mathcal{E}(A_j) + \sum_{j=1}^N \mu_A^2\right) = \\ &= \frac{1}{N}\left[N(\mu_A^2 + \sigma_A^2) - 2N\mu_A^2 + N\mu_A^2\right] = \\ &= \frac{\sigma_A^2}{N},\end{aligned}$$

kde sme využili definíciu estimátora (2.10), linearitu strednej hodnoty a výrazy (4.32) a (4.35).

Rovnako postupujeme v prípade strednej hodnoty estimátora $\check{\sigma}_A^2$

$$\begin{aligned}
\mathcal{E}(\check{\sigma}_A^2) &= \mathcal{E}\left(\frac{1}{N}\sum_{j=1}^N\left(A_j - \frac{1}{N}\sum_{k=1}^N A_k\right)^2\right) = \\
&= \frac{1}{N}\mathcal{E}\left(\sum_{j=1}^N\left[A_j^2 - \frac{2}{N}A_j\sum_{k=1}^N A_k + \left(\frac{1}{N}\sum_{k=1}^N A_k\right)^2\right]\right) = \\
&= \frac{1}{N}\left(\sum_{j=1}^N\mathcal{E}(A_j^2) - \frac{2}{N}\sum_{j=1}^N\sum_{k=1}^N\mathcal{E}(A_j A_k) + \frac{1}{N^2}\sum_{j=1}^N\sum_{k=1}^N\sum_{l=1}^N\mathcal{E}(A_k A_l)\right) = \\
&= \frac{1}{N}\left\{N(\mu_A^2 + \sigma_A^2) + \left(-\frac{2}{N} + \frac{1}{N^2}N\right)[N(\mu_A^2 + \sigma_A^2) + (N^2 - N)\mu_A^2]\right\} = \\
&= \frac{N-1}{N}\sigma_A^2,
\end{aligned}$$

kde sme využili definíciu estimátora (2.12), linearitu strednej hodnoty a výrazy (4.32), (4.34) a (4.34).

V prípade posledného estimátora $\check{\sigma}_A^2$ strednú hodnotu nebudeme rátať komplikovane, využijeme iba linearitu strednej hodnoty a to, že estimátory $\check{\sigma}_A^2$ a $\check{\sigma}_A^2$ sú vzájomne lineárne

$$\begin{aligned}
\mathcal{E}(\check{\sigma}_A^2) &= \mathcal{E}\left(\frac{1}{N-1}\sum_{j=1}^N(A_j - \tilde{A})^2\right) = \\
&= \mathcal{E}\left(\frac{N}{N-1}\check{\sigma}_A^2\right) = \frac{N}{N-1}\mathcal{E}(\check{\sigma}_A^2) = \sigma_A^2.
\end{aligned}$$

Literatúra

- [1] SKÁLA, Lubomír. *Úvod do kvantové mechaniky*. 1. vyd. Praha: Academia, 2005. ISBN 80-200-1316-4.
- [2] LEVINE, Iva N. *Quantum chemistry*. 5. vyd. London: Prentice hall, 2000. ISBN 0-13-685512-1.
- [3] KALOS, Malvin H., WHITLOCK, Paula A. *Monte Carlo Methods*. 2. vyd. Weinheim: WILEY-VCH Verlag GmbH & Co., 2008. ISBN 978-3-527-40760-6.
- [4] PRESS, William H., TEUKOLSKY, Saul A., VETTERLING, William T., FLANNERY, Brian P. *Numerical Recipes in C*. 2. vyd. Cambridge: Cambridge University Press, 1999. ISBN 0-521-43108-5.
- [5] FLYVBJERG, H., PETERSEN, HG. Error estimates on averages of correlated data. *Journal of chemical physics*. 1989, vol. 91(1), s. 461-466. ISSN 0021-9606.
- [6] CELESTE, R., DACOSTA, H.F.M., DASILVA, A.B.F., TRSIC, M. On the helium ground-state Hartree-Fock energy. *Chemical physics letters*. 1991, vol. 183, s. 31-33. ISSN: 0009-2614.
- [7] DRAKE, G.W.F. High-precision variational calculations for the 1s21S state of H- and the 1s21S,1s2s 1S and 1s2s 3s states of helium. *Nuclear instruments & methods in physics research section B-beam interactions with materials and atoms*. 1988, vol. 31, s. 7-13. ISSN 0168-583X.
- [8] BERGESON, S.D., *et al.*. Measurement of the He ground state lamb shift via the two-photon 1(1)S-2(1)S transition. *Physical review letters*. 1998, vol. 80, s. 3475-3478. ISSN: 0031-9007.
- [9] MORTON, D.C., WU, Q.X., DRAKE, G.W.F. Energy levels for the stable isotopes of atomic helium (He-4 I and He-3 I). *Canadian journal of physics*. 2006, vol. 84, s. 83-105. ISSN: 0008-4204.
- [10] AMOVILLI, C., MARCH, N.H., TALMAN, J.D. Generalizing Hartree-Fock theory for nonrelativistic atomic ions to yield a ground-state electron density satisfying cusp and asymptotic conditions. *Physical review A*. 2008, vol. 77. ISSN: 1050-2947.