

Diploma Thesis

Bachelor Thesis Simon König

submitted in satisfaction of the requirements for the degree of
Diplom-Ingenieur
of the TU Wien, Faculty of Civil Engineering

Diplomarbeit

Bachelor Arbeit Simon König

ausgeführt zum Zwecke der Erlangung des akademischen Grades eines
Diplom-Ingenieurs
eingereicht an der Technischen Universität Wien, Fakultät für Bauingenieurwesen
von

Bachelor Arbeit Simon König

Matr.Nr.: 01234567

unter der Anleitung von

Prof. Dipl.-Ing. Dr.techn. **forename surname**

Dipl.-Ing. Dr.-Ing. **forename surname**, BSc

Institut für YYY
Forschungsbereich
Technische Universität Wien
Karlsplatz 13/YYY, 1040 Wien, Österreich

Wien, im Januar 2018

Kurzfassung

Abstract

Inhaltsverzeichnis

| | | |
|----------|--|-----------|
| 1 | Einleitung | 7 |
| 1.1 | Einleitung, Section 1 | 7 |
| 2 | TheorieTeil von VielKörper SG zur DFT | 8 |
| 2.1 | Vielkörperproblematik der analytischen QM, Grundlegend für Ein-Körper Problem | 8 |
| 2.2 | Vielkörperproblematik, Diskussion der Hartree-Fock Methode als ähnlicher, aber doch verschiedener Ansatz + Born Oppenheimer Näherung als wichtige Beiträge zur DFT | 9 |
| 2.2.1 | Vereinfachungen hervorstreichen | 9 |
| 2.2.2 | Vielkörperproblem der QM | 10 |
| 2.3 | Ansatz der DFT mit allen Termen | 12 |
| 2.3.1 | Warum DFT? | 12 |
| 2.3.2 | Annahmen der DFT, Energie als Funktional der Dichte | 12 |
| 2.4 | Lösung der DFT | 14 |
| 3 | Uebergangsteil von DFT zu ML Gauss Regression | 15 |
| 3.1 | Motivation, warum ML | 15 |
| 3.1.1 | Pfad von der DFT zu ML Methode | 15 |
| 3.1.2 | Was wird genau verwendet | 15 |
| 3.2 | Die Gauß Regression | 16 |
| 3.2.1 | Vorraussetzungen für Gaussian Process | 16 |
| 3.2.2 | Bauen der Covarianz Matrix, also trainieren des Modells | 17 |
| 3.2.3 | Vorraussagen machen mit Modell | 17 |
| 3.3 | Modifikationen zur quantenmechanischen Anwendung | 18 |
| 3.3.1 | Totale Energien und Kräfte | 18 |
| 3.3.2 | Beschränkung auf nahe Umgebung - cutoff radius | 18 |
| 3.3.3 | Reduzierung der Trainingsets auf repräsentative Sets, ohne viel Mathematik | 19 |
| 3.3.4 | Beschreibung der Deskriptoren - Berücksichtigen der Invarianzen -Bispectrum | 19 |
| 3.3.5 | SOAP, als tatsächlich verwendete Basis, die als Argumente die Atompositionen nimmt | 20 |
| 3.4 | Verwendete Software, QUIP, nur kurzer Text Absatz um zu erklären, dass die beschriebenen Konzepte in QUIP umgesetzt sind | 20 |
| 4 | Implementing Gaussian Process train-predict-analyse | 21 |
| 4.1 | Talk about given problem, visualise hydrogen and Split up the task - training, prediction and analysis | 21 |
| 4.2 | Training | 21 |
| 4.2.1 | Get Training Data -> DFT -> XML file | 21 |
| 4.2.2 | Transform Training Data, xml to xyz with shell script | 21 |
| 4.2.3 | Split up Training and Test Data and briefly explain why | 21 |
| 4.2.4 | Perform gap fit on training data | 21 |
| 4.3 | Prediction | 21 |
| 4.3.1 | Make Prediction using QUIP | 21 |

| | | |
|----------|--|-----------|
| 4.3.2 | Prediction Plot, real values against predicted values | 21 |
| 4.4 | Analysis | 21 |
| 4.4.1 | Analyse prediction/real values graphs using RSME and R2 Score | 21 |
| 4.4.2 | Plot Dissipation(Bindungsenergie) Kurve and explain | 21 |
| 5 | Outlook and Conclusion | 22 |
| 5.1 | Why is ML useful to compute atomic energies -> saves computational cost . . . | 22 |
| 5.2 | Outlook - Possible use cases -> calculating energies for more complex structures using more complex descriptors | 22 |
| 5.3 | kritzelChapter | 23 |

Kapitel 1

Einleitung

1.1 Einleitung, Section 1

$$\pi = \frac{x}{2}$$

Kapitel 2

Theorie Teil von VielKörper SG zur DFT

2.1 Vielkörperproblematik der analytischen QM, Grundlegend für Ein-Körper Problem

Die Lösung der Schrödingergleichung stellt eines der fundamentalsten Probleme der Nicht-Klassischen Physik dar. In seiner einfachsten zeitunabhängigen Form wird sie angeschrieben als

$$TODO : write easy SG eq.$$

mit:

$$TODO : Define parameter of simple SG eq.$$

Die Schrödingergleichung ist als Eigenwertproblem definiert, somit ist ein Set an Basiswellenfunktionen zu finden welche für einen entsprechenden Hamilton-Operator eine Energie liefern die das System beschreibt. Die Wahl des Hamiltonoperators ist systemspezifisch hat aber im allgemeinen immer die Form:

$$TODO : H = T + V$$

Durch den Hamiltonian wird also eine Energie beschrieben, die immer aus der Kinetischen Energie *TODO:make italic T* des Systems und aus der Potenziellen Energie *TODO:make italic V* des Systems besteht.

Betrachtet man beispielsweise den Fall des einfachsten Atoms, des *TODO: Wassserstoff-Atoms*, lässt sich die Gleichung exakt lösen. Symmetrieüberlegungen führen zu einer Betrachtung der "Bewegung zweier Teilchen im kugelsymmetrischen Feld. Wie in der klassischen Mechanik lässt sich das 2-Körperproblem miteinander wechselwirkender Teilchen in der Quantenmechanik auf ein 1-Körperproblem reduzieren."¹. Da das Proton im Kern eine viel größere Masse hat als das Elektron *TODO: eminus*, wird die Bewegung des Protons vernachlässigt. Dies ist die Kernaussage der *Born-Oppenheimer-Näherung*, welche auch bei Vielteilchenproblemen vorteilhaft ist und auf die im Weiteren noch genauer eingegangen wird. Somit zerfällt die Wellenfunktion, in die freie Bewegung des Massenmittelpunktes von Proton und Elektron und in die Relativbewegung des Elektrons selbst. Die Bewegung des Massenmittelpunktes im freien Raum wird vernachlässigt und für das Elektron lässt sich der Hamiltonian in der Form

$$TODO : Hamiltonian Wasserstoff f für eminus$$

anschreiben.

Und die gesamte Schrödingergleichung in kartesischen Koordinaten

$$TODO : SG für Wasserstoff$$

mit den Termen:

¹Pitschmann 2020, S. 6.

TODO : TERME erklären von SG Wasserstoff kartsisch

Nach einer Transformation auf Kugelkoordinaten nutzt man die Kugelsymmetrie des Feldes aus. Die Erhaltung des Drehimpulses im kugelsymmetrischen Feld erlaubt eine Separation der Eigenfunktion in einen winkelabhängigen Teil und eine Radialkomponente.²

TODO : Separierte Wellenfunktion in radial und Winkelteil

Die Lösungen der Winkelkomponente der Wellenfunktion sind die Kugelfunktionen³

TODO : Kugelfunktionen anschreiben

und die Lösung des Radialteils ergeben sich zu:⁴

TODO : Radialfunktion anschreiben

Fügt man den Produktansatz wieder zusammen und löst mit dem Hamiltonian des Wasserstoffs Atom das Eigenwertproblem so erhält man für jedes Energieniveau n^2 -Eigenfunktionen. Jedes Energieniveau ist also n^2 -fach entartet. Konkret ergeben sich die diskreten Eigenzustände, abhängig von der Hauptquantenzahl n , zu:⁵

TODO : Diskrete Eigenzustände

Die Eigenfunktionen werden durch die Hauptquantenzahl n , die Nebenquantenzahl l , und die magnetische Quantenzahl m_l charakterisiert und bilden ein Orthogonalsystem.

TODO : zeige Eigenfunktionen sind Orthogonalsystem

Die realste und wichtigste Eigenschaft der Wellenfunktionen wird erst bei der Messung eines quantenmechanischen Zustandes messbar. Die Aufenthaltswahrscheinlichkeit das Elektron im Volumenelement dV anzutreffen ist gegeben durch:⁶

TODO : Formel für Aufenthaltswahrscheinlichkeit wie in Physik 4

2.2 Vielkörperproblematik, Diskussion der Hartree-Fock Methode als ähnlicher, aber doch verschiedener Ansatz + Born Oppenheimer Näherung als wichtige Beiträge zur DFT

2.2.1 Vereinfachungen hervorstreichen

Im Vorangegangenen Abschnitt wurde die Schrödingergleichung erfolgreich für ein sehr vereinfachtes System analytisch exakt gelöst und Energieeigenwerte E_n gefunden. Zur Simplizität des Systems kommen noch das Nicht-Berücksichtigen von experimentell festgestellten Phänomenen wie dem Spin oder Vakuumfluktuationen. Der Spin führt außerdem direkt ein magnetisches Moment ein, welches demzufolge auch außen vorgegeben wird. Weitere vereinfachende Annahmen sind:

²vgl. Zürich o.D., S. 218.

³vgl. Zürich o.D., S. 222.

⁴vgl. Zürich o.D., S. 230.

⁵vgl. Zürich o.D., S. 234.

⁶vgl. Zürich o.D., S. 234.

- Elektron wird als nicht relativistisches Teilchen betrachtet
- Wechselwirkung zwischen Elektron und Proton wird als Coulomb-Wechselwirkung angenommen
- Proton befindet sich in Ruhe

2.2.2 Vielkörperproblem der QM

Der bisher verwendete Hamiltonian und die zugehörigen Eigenfunktionen beschreiben ein sehr simples System. Reale Problemstellungen beziehen sich auf größere Systeme, sowohl von Elektronen als auch Kernen. Schon einfache Moleküle besitzen eine Vielzahl an Teilchen und sind mit analytischer Mathematik nicht mehr lösbar. Als Vorstufen zur Dichtefunktionaltheorie gilt es noch weitere Approximation zu berücksichtigen. Die im folgenden verwendeten Methoden finden sich in den Überlegungen zu DFT wieder.

Zwei notwendige Approximationen beziehungsweise Ansätze stellen die

- Born-Oppenheimer Näherung, und die
- Hartree- bzw. Hartree-Fock Methode

dar.

In atomare Einheiten kann der, ein System aus N_n Kernen und N_e Elektronen beschreibendem, Vielkörper Hamiltonian geschrieben werden als:⁷

TODO : schreibe vielkörper hamiltonian an, sowie wie hande ihn anschreibt

Mit der Einheitentransformation auf atomare Einheiten:

TODO : Einheitentransformation auf atomare Einheiten

2.2.2.1 Born-Oppenheimer Näherung

Betrachtet man den Operator der kinetischen Energie *TODO : Operator T* in einem Massenschwerpunktsystem mit einem Nukleus und N Elektronen, ähnlich dem des Wasserstoffatoms. Im Allgemeinen Fall gibt es nicht ein e^- sondern N Elektronen.

TODO : Operator der Kinetischen Energie ohne Born Oppenheimer Näherung

Born und Oppenheimer schlagen vor, aufgrund der großen Massendifferenz von Elektron und Proton, die Bewegung des Protons gegenüber der des Kernels zu vernachlässigen. Dem zugrunde liegt die Idee, dass sich Elektronen aufgrund ihrer um den Faktor 10^3 kleineren Masse viel schneller bewegen als die schweren Kerne. Veranschaulicht bedeutet das, die Elektronen sind bei jeder Bewegung der Kerne schon wieder in ihren Grundzustand relaxiert. Aufgrund dessen kann die Bewegung des Kernels vernachlässigt werden und gehen nicht mehr in den Hamiltonian mit ein. Somit vereinfacht sich der Operator der kinetischen Energie im Limes *TODO lim M gegen unendlich*.⁸

TODO : Operator der Kinetischen Energie MIT Born Oppenheimer Näherung

Die Born-Oppenheimer Näherung wird eingeführt für ein System aus einem Nukleus und N Elektronen. Ihr Konzept ist umlegbar auf einen allgemeinen Hamiltonian mit N_n Kernen und

⁷Lecture 4 Hande o.D., S. 2.

⁸Hjorth-Jensen 2008, S. 272.

N_e Elektronen gemäß TODO:gleichungsreferenz. Die Born-Oppenheimer Näherung geht in die Dichtefunktionaltheorie instrumental ein.

2.2.2.2 Hartree-bzw. Hartree Fock Methode

Die Hartree Methode, und im weiterführenden die Hartree Fock Methode macht einen Ansatz für die Vielelektron Wellenfunktionen. Die Terme des allgemeinen Vielteilchenhamiltonians aus (TODO: Gleichung für VielTeil.Hamiltonian) welche sich auf Kern-Kern Wechselwirkung beziehen werden als konstant angenommen. Diese Annahme ist durch die Born-Oppenheimer Näherung begründet.

Der Elektronen-Hamilton Operator lässt sich schreiben als:

TODO : schreibe elektronenhamiltonian an, sowie wie hande ihn anschreibt

Die ersten beiden Terme sind Summen über Ein-Teilchen Operatoren und beziehen sich nur auf eine Elektron-Koordinate. Diese Beiträge zur Energie sind trivial, da sie äquivalent sind zu Lösungen der Einteilchen-Schrödingergleichungen für die jeweiligen Elektronen, anschließend wird über diese einzelnen Energiebeiträge summiert. Die Schrödingergleichung separiert hier, weil jeder Ein-Teilchen Operator nur eine Wellenfunktion "*sieht*", welche nur abhängig von der Koordinate des jeweiligen Teilchens ist.⁹

Der interessante Term ist demfolgend:

TODO : schreibe elektronenhamiltonian an, sowie wie hande ihn anschreibt aber nur den letzten Term über die zwei

Aus dieser Gleichung liest sich das Potential für das i -te Elektron, hier wieder mit dem Elektron-Kern Potential, wie folgt:

Potential für i – tes Atom wie auf Seite 77 in akt1 skript

Hier wird ein Problem sichtbar, welches ebenso in der Dichtefunktionaltheorie auftaucht. Das Potential im Hamiltonian, dass ja zur Bestimmung der Eigenfunktionen notwendig ist, hängt von ebenen jenen Eigenfunktionen selbst ab. Die Näherung von Hartree kann als "Zentralfeld-näherung" bezeichnet werden, da sie jeweils das Potential des i -te Elektrons als Summe über die Wahrscheinlichkeitsdichten aller anderen Elektronen modelliert. Eine ähnliche Überlegung findet sich in der DFT wieder.¹⁰

Hartree verwendet als Lösungsansatz das direkte Produkt der Elektronenorbitale. Dieser Ansatz berücksichtigt das Pauli Prinzip nicht.

Hartree Ansatz für Einteilchen – Produkt Wellenfunktion

Da nun Elektronen Fermionen sind und deshalb bei einer Symmetrievertauschung, formal der Anwendung des Paritätsoperators, einer Vorzeichenänderung unterliegen, genügt dieser einfache Produktansatz nicht. Eine Verbesserung legen *Hartree & Fock* dar, und zwar die Ersetzung der Orbitalprodukte durch Slaterdeterminanten. Die Darstellung der Wellenfunktion durch Slaterdeterminanten welche der Antisymmetrie-Bedingung genügt erfüllt das Ein-Teilchen Bild.¹¹

Write Slater – Determinant like hande

⁹Lecture 4 Hande o.D., S. 1.

¹⁰Pitschmann 2020, S. 78.

¹¹Lecture 4 Hande o.D., S. 2.

Die Lösung ist aufgrund der direkten Abhängigkeit des Hamiltonians von den durch ihn zu bestimmenden Eigenfunktionen auch für eine günstige Wahl der Eigenfunktionen nicht eindeutig lösbar. Eine häufig verwendete Lösungsmethode ist die des *Selbst-Konsistenten Feldes*. Die numerische Lösung setzt einen gut gewählten ersten Ansatz der Wellenfunktionen voraus und optimiert diese dann gemäß einer iterativen Methode:

1. Ansatz für Eigenfunktionen, verschiedene Methoden möglich (z.B.: Thomas-Fermi-Gleichung)
2. Potentiale werden bestimmt mittels Aufenthaltswahrscheinlichkeit: *TODO* Aufenthaltswahrscheinlichkeit
3. Neue Eigenfunktionen werden mit dem gegebenen Potential bestimmt

Der Vorgang wird iterativ numerisch gelöst bis sich eine *selbstkonsistente* Lösung einstellt.

2.3 Ansatz der DFT mit allen Termen

2.3.1 Warum DFT?

Die Lösung der Schrödingergleichung ist durch die Approximationen von Born-Oppenheimer beziehungsweise Hartree-Fock einfacher geworden. Hartree-Fock separieren die Eigenfunktionen eines N-Teilchen Systems in ein Produkt aus N-Orbitalen für jeweils eine Raumkoordinate. Dieser Ansatz braucht für ein Elektron nun 3 Raumdimensionen, für zwei Elektronen 6 Raumkoordinaten und für N Elektronen $3 \cdot N$ Raumkoordinaten. Dieser Ansatz bringt bereits eine Reduktion des Rechenaufwandes im Gegensatz zu den 3^N Raumdimensionen, die für eine exakte Lösung der N-dimensionalen Schrödingergleichung notwendig sind. Für reale Vielkörperprobleme sind immer noch Kalkulationen von extremem Aufwand notwendig. Ein simples CO_2 Atom besitzt 22 Elektronen (*TODO*: Grafik einfügen), die Schrödingergleichung wird ein 66-dimensionales Problem.

$$\text{TODO} : \text{CO}_2 \text{Atomanschriften} = 22 \text{elektronenanschriften}$$

Die Konfiguration eines Blei Einheitszelle, mit etwa 100 Atomen, ergibt bei 14 Elektronen pro Pb-Atom, ein 42 000-dimensionales Problem.

$$\text{TODO} : \text{SiliziumEinheitszelleAnschriften} = 42000 \text{elektronen}$$

Mit Quanten-Chemie Methoden wie zum Beispiel der *Møller-Plesset perturbation theory* beträgt die Rechendauer um ein derartiges Problem, wie das der Blei Einheitszelle, zu lösen, zirka ein Jahr. Mithilfe der *Dichtefunktionaltheorie* ist eine Lösung in etwa 5 Stunden realistisch. (*TODO*: zitiere <https://ocw.mit.edu/courses/materials-science-and-engineering/3-021j-introduction-to-modeling-and-simulation-spring-2012/part-ii-lectures-videos-and-notes/lecture-3/>)

2.3.2 Annahmen der DFT, Energie als Funktional der Dichte

Walter Kohn schreibt in seinem Artikel, zu Ehren seiner Nobelpreisverleihung, die der Dichtfunktionaltheorie zugrundeliegende Aussage:¹²

"The basic lemma of HK. The ground-state density $n(\mathbf{r})$ of a bound system of interacting electrons in some external potential $v(\mathbf{r})$ determines this potential uniquely (Hohenberg and Kohn, 1964)"

¹²Kohn 1999, S. 7.

Ist die Energie eines Systems nur von der Elektronendichte abhängig, bedeutet das, für eine beliebig große Elektronenkonfiguration, eine Reduktion auf 3 Raumkoordinaten. Der Hamilton-Operator eines Systems, in der Born-Oppenheimer-Näherung, lässt sich schreiben als:

TODO : Schreibe $H = T + V_{ext} + V$ mit Dichte als n wie in FK2

TODO : beschreibe diese Terme wie in Hutter

Die Theorie von Kohn, Hohenberg und Sham geht auf das Jahr 1964 zurück und fußt auf zwei fundamentalen Theoremen:

1. V_{ext} ist bis auf eine Konstante eindeutig als Funktional der Elektronendichte *TODO : n bestimmt.*
2. Das Variationsprinzip belegt dieses Theorem eindeutig. (TODO: maybe Beweis skizzieren)
3. Das Funktional V_{ext} hat sein Minimum gegenüber einer Variation *TODO : variation – teilchendichte* der Teilchendichte für eine Gleichgewichtsdichte n_0 für gegebenes V_{ext} .
4. Das zweite Theorem ist ebenfalls eine Folge des allgemeinen Variationsprinzips der Quantenmechanik (TODO: maybe Beweis skizzieren)

Allgemein kann die Energie auf Basis der Hohenberg-Kohn-Shams-Theoreme (HKS) für ein System aus N wechselwirkenden Teilchen im externen Potential V_{ext} geschrieben werden als¹³

TODO : schreibe $E_{vonn} = 4$ Terme wie bei dft – hutter seite 6 ganz unten

TODO : Erkläre Terme

Zunächst betrachtet man den einfacheren Fall von N nicht-wechselwirkende Teilchen. Der Hamiltonian ist gegeben als:

TODO : Hamiltonian ohne V_e Wechselwirkung

Die Eigenfunktionen der Schrödinger-Gleichung

TODO : SG für nicht – wechselwirkendes – System

in einem Nicht-Wechselwirkenden-System können mittels Produktansatz gelöst werden und sind, aufgrund der Natur der Elektronen als Fermionen, durch die Slater-Determinanten der einzelnen Orbitale gegeben. Die Elektronendichte ist dann:

TODO : elektronendichte anschreiben als summe über eigenfunktionen

Die Kinetische Energie T_s als Funktional der Dichte *TODO elektronendichten* ist dann:

TODO : Funktional von T wie dft – hutter seite 6

Im allgemeinen Fall, wie in der Gleichung (TODO: reference zu E von $n = 4$) beschrieben, wechselwirken die Elektronen untereinander. Addiert und subtrahiert man das bereits bestimmte T_s von der Gleichung (TODO: reference zu E von $n = 4$) ergibt sich¹⁴

¹³Hutter 2004, S. 6.

¹⁴Hutter 2004, S. 7.

TODO : Gleichung für E_0 aber mit E_{XC} wird ft – hutterseite 7 oben

mit dem Austausch-Korrelationspotential:

TODO : Gleichung für E_{XC} wird ft – hutterseite 7 fast oben

Die Grundzustandsenergie als Funktional der Elektronendichte besteht nun aus lösbaren Termen und dem Austausch-Korrelationspotential E_{XC} . Um die Dichte zu kalkulieren werden die Kohn-Sham-Funktionen benötigt, diese sind Eigenfunktionen der Kohn-Sham-Gleichungen¹⁵

TODO : Kohn – Sham – Gleichung wie in FK2 Seite 161 oben, mit $V = V_{ext}$

mit

TODO : $\mu = \text{Variation von } E_{XC}, \text{ FK2 Seite 160}$

Die Kohn-Sham-Gleichungen sind verallgemeinerte Ein-Teilchen-Schrödingergleichung in denen ein effektives Potential V_{eff} anstatt eines Vielteilchenpotentials angeschrieben wird. Sie berücksichtigen das Austausch-Korrelationspotential E_{XC} .

Die Lagrange-Multiplikatoren ϵ_i sind orthogonalitätserhaltend für die Kohn-Sham-Orbitale. Bei der Kalkulation der Eigenfunktionen der Kohn-Sham Gleichung tritt ein ähnliches Problem zu dem der Hartree-Fock Methode auf: Der Kohn-Sham-Hamiltonian ist direkt von der Elektronendichte abhängig. Zur numerischen Lösung der Kohn-Sham-Gleichungen und somit zum finden der Elektronendichte kann die Methode des Selbst-Konsistenten-Feldes, wie bei Hartree-Fock, verwendet werden.

Die Kohn-Sham-Gleichungen mit ihren iterativen Lösungsmethoden stellen die Grundlage von *ab-initio* Methoden dar, zu denen auch die in dieser Arbeit verwendete Methode VASP (Vienna Ab Initio Package) gehört.¹⁶

2.4 Lösung der DFT

¹⁵Paschen und Mohn 2020, S. 161.

¹⁶Paschen und Mohn 2020, S. 161.

Kapitel 3

Uebergangsteil von DFT zu ML Gauss Regression

3.1 Motivation, warum ML

3.1.1 Pfad von der DFT zu ML Methode

Die Ermittlung von Energien atomarer Systeme mittels Dichtefunktionaltheorie bringt eine Steigerung gegenüber der exakten Lösung der Schrödingergleichung. Im Allgemeinen Fall, sind quantenmechanische Systeme allerdings so komplex, dass auch hier der Rechenaufwand den realistisch Möglichen übersteigt. Weiters beschreiben die Simulationen durch die DFT ein genaues Setting und sind durch ihre empirisch berechneten Funktionale limitiert. Die Motivation *Machine-Learning* Methoden einzusetzen ist die genauen Beschreibung einer *Born-Oppenheimer Fläche der potentiellen Energie (PES)*.

Bisher ist das Konzept von *PES* in Verwendung, um Energieflächen einer fixierten Atom-Konstellation zu modellieren. Mittels Machine-Learning Algorithmen wird versucht eine klare Energiefläche zu finden, welche die Energien einzelner Atome oder Atom-Bindungen, als Funktion ihrer Umgebung (*neighbourhood-environment*) beschreibt. Ein solches System reagiert empfindlich auf Änderungen seiner Umgebung, wie zum Beispiel eine Veränderung der Anzahl, oder des Typs, der Umgebungs-Atome.¹

Jeder, und auch die hier besprochenen Machine-Learning-Algorithmen, sind nicht per se in der Lage atomare Systeme zu Lösen und Energien zu berechnen. Sie beruhen auf Trainingsdaten und Testdaten und benötigen daher eine vorausgehende quantenmechanische Methode um Daten-Sets zu generieren. Im Allgemeinen gilt, dass ML- und insbesondere die hier verwendeten Gauß-Regressions-Modelle, sehr gut für Interpolation geeignet sind, nicht jedoch für Extrapolation. Das heißt, das trainierte *Machine-Learning*-Modell funktioniert gut in den Bereichen der Trainingsdaten, nicht jedoch für exotischere Atom-Konstellationen, die nicht trainiert wurden.

3.1.2 Was wird genau verwendet

In dieser Arbeit wird zur Energieberechnung von atomaren Systemen ein, sich stochastischen Prozessen bedienendes, *Framework* verwendet. Eine andere Möglichkeit wäre die Verwendung einfacher *Neuronaler Netzwerke (NN)*. Es zeigt sich, dass ein geeignetes stochastisches Werkzeug der Gauß'sche Regressions-Prozess ist.

Um die mittel Regression Vorhersagen treffen zu können wird die atomare Energie eines Sets von Atomen, als Funktion ihrer Atom-Konfiguration, das heißt ihrer Geometrie eingeführt. Diese Energie besteht aus der Summe über alle Energiefunktionale innerhalb des "Cut-Off-Radius" und einem vernachlässigbaren (long-range-) Term. Die Beiträge des letzten Terms beschreiben elektrostatische Kräfte, beispielsweise Polarisierung oder Van-der-Waals Kräfte.

¹P. und G. 2014, S. 1051.

TODO : $E = \text{DOPPELSUMME}_{\alpha} \text{wieinGAP} - \text{Intro, Seite 1051}$

Die Summe über α erstreckt sich über die verschiedenen Typen an Energiefunktionalen. Das sind im einfachsten Fall Zwei- oder Drei-Körper Energien, die sich durch atomare Abstände, oder Winkel zwischen den Atomen beschreiben lassen. Die Argumente dieser Energiefunktionale nennt man Deskriptoren. Deskriptoren beschreiben die Atomkonfiguration, mittels Transformation der kartesischen Koordinaten der Atome in Umgebungskoordinaten für die gesuchte lokale Energie eines Ziel-Atoms. Die Summe über i läuft über alle Instanzen von Bindungs-Winkeln respektive Bindungs-Abständen.²

3.2 Die Gauß Regression

3.2.1 Voraussetzungen für Gaussian Process

Der Gauß'sche Prozess kann zur Lösung von Regressionsproblemen eingesetzt werden. Dazu nimmt man ein fixiertes Set von Basisfunktionen \mathbf{H} an, in deren Raum der Dateninput \mathbf{X} lebt. Die Dimension von X sei N und die Matrix \mathbf{R} definiert den Input *TODO : $x_{hoch}(n)$* in der Basis \mathbf{H} . Für die Wahl der Basis ergeben sich je nach Problemstellung unterschiedliche Optionen. Einfach Wahlen wären beispielsweise radiale oder polynomiale Basisfunktionen. Umgelegt auf die Auffindung von quantenmechanischen Energien, wird das Input-Datenset von Deskriptoren \mathbf{d}_N beschrieben. Die Deskriptoren stehen für die betrachtete atomare Umgebung und sollen klare Vorhersagen über die zugehörigen Energien machen.

TODO : Formel 45.17 aus GAP – Infotheory

Die vorhergesagten Werte werden mit \mathbf{y}_N angenommen und sind auch von Dimension N . Der Input-Datensatz \mathbf{X} in der Basis \mathbf{H} wird durch einen Gewichtsvektor \mathbf{w} auf die Ergebnisse \mathbf{y}_N gemapped. Der Vektor \mathbf{w} hat Rang N .

TODO : Formel 45.18 aus GAP – Infotheory

Mit *David J.C. MacKay* nimmt man die Wahrscheinlichkeitsverteilung der vorhergesagten Werte y als Normalverteilung an:

TODO : Formel 45.19 aus GAP – Infotheory

und die hierbei verwendete Kovarianzmatrix von \mathbf{y} ist folgendermaßen definiert:³

TODO : Formel 45.20 aus GAP – Infotheory

Die Hauptannahme hier ist, dass die Verteilung der Gewichte als Gauß'sche Normalverteilung angenommen wird. Dadurch fällt die, im Normalfall schwierige, *fitting*-Aufgabe, dass auffinden der Gewichte weg und wird mithilfe der Varianz *TODO : $\sigma^2_{hoch2undindexw}$* modelliert. Im Weiteren wird somit die Kovarianzmatrix, auch *Kernel* genannt, das tragende Prognose-Instrument.⁴

²P. und G. 2014, S. 1051.

³MacKay 2003, S. 540.

⁴P. und G. 2014, S. 1052.

3.2.2 Bauen der Kovarianz Matrix, also trainieren des Modells

Jedes Datenset \mathbf{X} in einer gewählten Basis \mathbf{H} induziert also ein Kovarianzmatrix, und nur dieses *Kernel* wird für weitere Vorhersagen benötigt. Das Modell muss noch *trainiert* werden. Dazu benötigt man *target values* \mathbf{t} . In dieser Arbeit sind diese Werte die Energieniveaus der, mit den zugehörigen Deskriptoren berechneten, quantenmechanischen Systeme. Die Energien werden aus der Dichtefunktionaltheorie berechnet. Die Dimension von \mathbf{t} ist N . Das Regressionsmodell nimmt also ein Set an Deskriptoren $TODO : R(d_n)$ und trainiert es mit den entsprechenden Energiewerten \mathbf{t}_N .⁵

Aufgrund der Gauß'schen Modellierung, muss angenommen werden, dass jedes der *target values* \mathbf{t} auch um ein Gauß'sches vom vorhergesagten Wert \mathbf{y} abweicht. Das führt dazu, dass \mathbf{t} nicht scharf definiert ist, sondern auch einer Normalverteilung unterliegt.

TODO : Formel 45.23 aus GAP – Infotheory

Das führt zu einer *trainierten* Kovarianzmatrix mit vollem Rang:⁶

TODO : Formel 45.24 aus GAP – Infotheory

Anzumerken ist, dass die Wahl der Kovarianzmatrix, sprich des *Kernels*, im Allgemeinen nicht vorbestimmt ist. Die einzige Einschränkung ist, dass die Matrix nicht negativ definit werden darf. Das muss für jedes Set \mathbf{X}_n an Datenpunkten gelten.

3.2.3 Vorraussagen machen mit Modell

Die eigentliche Aufgabe des Modells ist es Vorhersagen zu machen, nachdem es mit dem Datenset, bestehend aus \mathbf{x}_N und \mathbf{t}_N trainiert wurde. Der $N+1$ wert für \mathbf{t} ist gesucht, also $t(N+1)$.

Die Kovarianzmatrix \mathbf{C} wird also um einen Eintrag erweitert und ist nun in Dimension $N+1$. Zur einfacheren Darstellung, definiert man $\mathbf{C}_{(N+1)}$ neu:⁷

TODO : Formel 45.35 aus GAP – Infotheory

Die Wahrscheinlichkeitsverteilung, unter Berücksichtigung eines neuen, unbekannten Wertes, ist Gauß-verteilt und benutzt als *Kernel* $\mathbf{C}_{(N+1)}$:

TODO : Formel 45.36 aus GAP – Infotheory

Als vorhergesagten Wert $t(N+1)$ nimmt man nun den Mittelwert dieser Verteilung an:

TODO : Formel 45.42 aus GAP – Infotheory

Hier zeigt sich die Einfachheit des Gauß'schen Modellierungs Prozesses. Zur Prognose neuer Energiewerte, wird nicht die Berechnung der Matrix $\mathbf{C}_{(N+1)}$ benötigt, sondern nur die Invertierung der Matrix \mathbf{C}_N . Somit sind einzig die, durch die Basis mit den Deskriptoren \mathbf{d}_N induzierte, Kovarianzmatrix \mathbf{C}_N und die Trainings-Energieniveaus \mathbf{t}_N notwendig, um Interpolation durchzuführen.⁸

⁵Bartók u. a. 2009, S. 3.

⁶MacKay 2003, S. 540.

⁷MacKay 2003, S. 543.

⁸P. und G. 2014, S. 1052.

3.3 Modifikationen zur quantenmechanischen Anwendung

3.3.1 Totale Energien und Kräfte

Die Dichtefunktionaltheorie, sowie auch andere quantenmechanische Berechnungsmethoden, ermitteln Energien kompletter Systeme für eine gegebene Konfiguration. Der erklärte Gauß'sche Approximations Prozess nimmt lokale atomare Energien und Atom-Geometrien als Argumente. Lokale Energien und lokale Gradienten werden aus den quantenmechanisch berechneten Totalen Energien und ihre Totalen Gradienten abgeleitet. Die Berechnung der Kovarianz erfolgt analog zu der für lokale Energien mit zusätzlicher Summation über alle Atome des Systems. Die Analogie zur lokalen Kovarianz-Matrix ist gegeben.⁹

TODO : Formel12 aus GAP – Intro, bisserl kürzer und paar Schritte auslassen

Die Kraft eines Systems ist durch die Ableitung nach dem Ort gegeben. Die Kraft im Bezug auf die Totale Energie ist gegeben als Gradient nach allen Atomen k nach allen kartesischen Koordinaten α .

TODO : Formel13 aus GAP – Intro, bisserl andere Notation

Die Gradienten werden nach den Deskriptoren gebildet und sind je nach Wahl der Deskriptoren zu berechnen. Im Falle der Ableitungen der Kovarianzmatrix ergibt sich eine Jacobi-Matrix durch Ableitung nach Koordinaten in beiden Systemen.¹⁰

TODO : Formel17 aus GAP – Intro, bisserl andere Notation

TODO : Wo bei elementar aus N und j elementar aus M , umschreiben in mathematische Schreibweise

3.3.2 Beschränkung auf nahe Umgebung - cutoff radius

Die Berechnung lokaler Energien beschränkt sich auf eine gewisse Umgebung der betrachteten Atome. Die Größe der Umgebung ist nicht eindeutig definiert und wird durch die gewünschte Genauigkeit der Ergebnisse induziert. Wird ein großes Umfeld mit einbezogen steigt der Rechenaufwand stark an, da mit der Größe des Umfelds die Anzahl der betrachteten Atome steigt. Hingegen liefert eine zu klein gewählte Umgebung ungenaue und nicht repräsentative Ergebnisse.

Ein, die Umgebung begrenzender, Radius muss gewählt. Die lokale Energie geht fließend Richtung Null, sobald der radiale Abstand des betrachteten Atoms den Begrenzungs-Radius überschreitet. Eine einfache Veranschaulichung für eine beliebige lokale Energie:^{11,12}

TODO : Formel22 aus GAP – Intro

Der Begrenzungs-Radius hat die Form:¹³

TODO : Formel22 aus GAP – Intro, nur als mittlere Funktion eine allgemeine Funktion von r_c und d angeben

⁹P. und G. 2014, S. 1053.

¹⁰P. und G. 2014, S. 1053.

¹¹P. und G. 2014, S. 1053.

¹²Bartók u. a. 2009, S. 2.

¹³P. und G. 2014, S. 1054.

Hier ist *TODO* : *allg.Funktion von* (r, r_{begr}, d) eine Funktion die von r , r_{begr} und d abhängt, mit d als veränderlicher Parameter. Der Parameter d determiniert die Größe der Übergangsregion in der Nähe des Cutoffradius. Die Parameter d und r_{begr} können bei einer möglichen Applikation vom Anwender definiert werden.

Diese Begrenzung auf eine bestimmte lokale Umgebung kann entweder in die Bildung der Kovarianz-Funktion oder direkt in die Deskriptoren implementiert werden.

3.3.3 Reduzierung der Trainingsets auf repräsentative Sets, ohne viel Mathematik

Die Berechnung der Kovarianz Matrix und deren Jacobi-Matrizen ist aufwendig und kann vereinfacht werden. Im Bereich des *Maschinellen Lernens (ML)* ist der Praktik der *Ausdünnung* (engl.: Sparsification) verbreitet. Die Anzahl der Atom-Konfigurationen und die dazugehörigen lokalen Energien werden verkleinert, die Trainingsdaten werden also minimiert. Dazu muss, beispielsweise durch Zufallsprinzip oder anderen ML-Methoden, eine repräsentative Menge an Atom-Konfigurationen gewählt. Um Genauigkeit zu gewährleisten werden die lokalen Energien der Trainingsdaten zu einer kleineren Menge an Punkten linear-kombiniert. Die Berechnung der lokalen atomaren Energien unter Verwendung von Ausdünnung im Folgenden:¹⁴¹⁵

TODO : 2.3 aus Anhang von Gap-2009 plus Erklärung aller Terme gemäß, 2. aus Anhang. Viele Dicke Formeln hier

3.3.4 Beschreibung der Deskriptoren - Berücksichtigen der Invarianzen - Bispectrum

Deskriptoren beschreiben die lokale Anordnung von Atomen. Sie formen die Geometrie und durch sie wird die lokale Energie determiniert. Haben Deskriptoren zweier Konfigurationen ähnliche Werte, dann haben sie auch ähnliche Energie. Liegt eine spezifische atomare Anordnung vor, also Deskriptoren mit bestimmten Werten, dann wird mit dieser atomaren Anordnung, unter Verwendung quantenmechanischer Methoden (beispielsweise der DFT), die zugehörige Energie berechnet. Diese Paarung, also Geometrie und Energie sind die Daten, mit denen der Gauß'sche Approximierungs Prozess trainiert wird.

Die einfachsten Deskriptoren sind inter-atomare Abstände und Winkel. Für ein Zwei-atomiges Molekül dient der skalare Abstand als Deskriptor. Ein drei-atomiges Molekül benötigt als Deskriptor zumindest drei inter-atomare Abstände oder zumindest einen inter-atomaren Winkel. Die Komplexität nimmt mit der Größe der Systeme zu.

Die Energie eines atomaren Umfelds besitzt bestimmte Invarianzen. Sie ist invariant gegenüber Permutation, Translation und Rotation der Atome. Diese Symmetrien müssen also in der Erstellung des Modells berücksichtigt werden. Die Invarianz des Modells ist nicht automatisch gegeben und muss entweder grundlegend in die Deskriptoren eingebaut, oder die Kovarianz-Matrix im Weiteren modifiziert werden.

Die erste Möglichkeit ist die Implementierung auf Deskriptor-Ebene. Deskriptoren im kartesischen Raum berücksichtigen die notwendigen Symmetrien ebenso wenig wie dreidimensionale sphärische Koordinaten. Eine Möglichkeit um eine genau Beschreibung des Systems, unter Berücksichtigung aller Invarianzen zu gewährleisten ist die Projektion der atomaren Dichte auf eine Kugel im 4-dimensionalen Raum.¹⁶

¹⁴P. und G. 2014, S. 1054.

¹⁵Bartók u. a. 2009, S. 3.

¹⁶Bartók u. a. 2009, S. 3.

Projektion von 3d Koordinaten auf 4dsphäre, -- > bispectrum.eventuell neue Quelle

Die zweite Möglichkeit, Invarianzen der Energie zu berücksichtigen, ist eine Modifikation der Kovarianz-Matrix. Bei der Bildung des Kernels werden entsprechend Rotations-, Translations-, oder Permutationen auf die Basis-Funktionen angewandt. Jeder Satz an Basisfunktionen induziert eine Kovarianzmatrix. Die Basisfunktionen nehmen die Deskriptoren als Argumente. Das invariante Kernel muss normalisiert werden um die Skalierung der Ergebnisse nicht zu verändern. Um Invarianz gegenüber Rotation zu gewährleisten ist die Anwendung eines Rotations-Operators beim berechnen der Kovarianz-Matrix notwendig:¹⁷

GAP-intro Formel 40 und 41. Anwendung von R^u m Rotationsinvarianz zu gewährleisten, erklären von p .

3.3.5 SOAP, als tatsächlich verwendete Basis, die als Argumente die Atompositionen nimmt

VIELLEICHT, weil möglicherweise werden andere Deskriptoren verwendet in der Arbeit.

3.4 Verwendete Software, QUIP, nur kurzer Text Absatz um zu erklären, dass die beschriebenen Konzepte in QUIP umgesetzt sind

¹⁷P. und G. 2014, S. 1055.

Kapitel 4

Implementing Gaussian Process train-predict-analyse

4.1 Talk about given problem, visualise hydrogen and Split up the task - training, prediction and analysis

4.2 Training

4.2.1 Get Training Data → DFT → XML file

4.2.2 Transform Training Data, xml to xyz with shell script

4.2.3 Split up Training and Test Data and briefly explain why

4.2.4 Perform gap fit on training data

4.3 Prediction

4.3.1 Make Prediction using QUIP

4.3.2 Prediction Plot, real values against predicted values

4.4 Analysis

4.4.1 Analyse prediction/real values graphs using RSME and R2 Score

4.4.2 Plot Dissipation(Bindungsenergie) Kurve and explain

Kapitel 5

Outlook and Conclusion

- 5.1 Why is ML useful to compute atomic energies → saves computational cost**
- 5.2 Outlook - Possible use cases → calculating energies for more complex structures using more complex descriptors**

5.3 kritzelChapter

hjorth-jensen-try¹ toulouse-try² intro-ml-try³ hands-on-ml-try⁴ esl2-try⁵

¹Hjorth-Jensen 2008, S. 500.

²Toulouse 2019, S. 301.

³Müller und Guido 2017, S. 301.

⁴Géron 2017, S. 301.

⁵Trevor Hastie 2008, S. 301.

Literatur

- Kohn, W. (1999). *Nobel Lecture: Electronic structure of matter—wave functions and density functionals*.
- MacKay, David J.C. (2003). *Information Theory, Inference, and Learning Algorithms*. Cambridge University Press.
- Hutter, Jürg (2004). *Dichtefunktionaltheorie*.
- Hjorth-Jensen, Morten (2008). *Computational Physics*.
- Trevor Hastie Robert Tibshirani, Jerome Friedman (2008). *The Elements of Statistical Learning, Second Edition*. Springer Series in Statistics. Springer.
- Bartók, Albert P. u. a. (2009). „Gaussian approximation potentials: The accuracy of quantum mechanics, without the electrons“. In:
- P., Bartók A. und Scány G. (2014). „Gaussian Approximation Potentials: A Brief Tutorial Introduction“. In: *J. Quantum Chem*, 115, 1051–1057. DOI: 10.1002/qua.24927.
- Géron, Aurélien (2017). *Hands-On Machine Learning with Scikit-Learn and TensorFlow*. O'Reilly Media, Inc. 1005 Gravenstein Highway North, Sebastopol, CA 95472. ISBN: 978-1-491-96229-9.
- Müller, Andreas C. und Sarah Guido (2017). *Introduction to Machine Learning with Python*. O'Reilly Media, Inc. 1005 Gravenstein Highway North, Sebastopol, CA 95472. ISBN: 978-1-449-36941-5.
- Toulouse, Julien (2019). *Introduction to density-functional theory*.
- Paschen, Silke Bühler und Peter Mohn (2020). *Festkörperphysik II*.
- Pitschmann, Mario (2020). *Atom-, Kern-, und Teilchenphysik I*.
- Hande, Toffoli (o.D.). *Lecture Notes on Principles of Density Functional Theory*. URL: <http://www.physics.metu.edu.tr/~hande/teaching/741.html>. accessed: 05.12.2020.
- Zürich, Quantum Device Lab ETH (o.D.). *Physik IV, Einführung in die Quantenmechanik*.