

修士論文

三次元畳み込みニューラルネットワーク による交換相関汎関数の開発

指導教員 伊庭斉志 教授

平成 31 年 1 月 31 日提出

情報理工学系研究科電子情報学専攻
48-176448 堀場 貴裕

要旨

密度汎関数法は誕生以来様々な物性の解析に適用され大きな成功を収めてきた。しかしバンドギャップなどの計算ではしばしば誤った結果を与える。その誤差は密度汎関数法における交換相関汎関数の近似精度の不十分さに起因している。本研究では既存の交換相関汎関数を3DCNNによりモデリングし、誤差を改善するようにパラメータを更新することでより高精度な汎関数の開発を試みた。

目次

論文要旨	i
第1章 序論	1
1.1 研究背景	1
1.2 本論文の構成	2
第2章 物質中の電子	3
2.1 断熱近似	3
2.2 一電子近似	4
2.2.1 ハートリー法	5
2.2.2 ハートリーフォック法	6
2.3 電子相関	8
2.3.1 電子相関を取り込むための手法	9
第3章 密度汎関数法	11
3.1 トーマスフェルミ理論	11
3.2 ホーエンベルグコーンの定理	12
3.2.1 ホーエンベルグコーンの第一定理	12
3.2.2 ホーエンベルグコーンの第二定理	14
3.3 コーンシャム方程式	15
3.4 コーンシャム方程式の解法	17
3.4.1 自己無撞着計算	18
3.5 密度汎関数法の拡張	19
第4章 交換相関汎関数	20
4.1 交換相関汎関数の分類	20
4.1.1 局所密度近似(LDA)汎関数	20

4.1.2	一般化勾配近似(GGA)汎関数	22
4.1.3	meta GGA汎関数	24
4.1.4	混成汎関数	24
4.1.5	半経験的汎関数	25
4.2	交換相関汎関数に対する要請	25
第5章	バンドギャップ	28
5.1	ヤナックの定理と $N \pm 1$ 電子系	29
5.2	自己相互作用	30
5.3	ギャップ誤差の改善法	31
5.3.1	混成汎関数	32
5.3.2	GW近似	32
5.3.3	DFT+U	33
第6章	ニューラルネットワーク	35
6.1	ニューラルネットワークの基本構造	35
6.2	誤差逆伝播法	37
6.2.1	ヤコビ行列	38
6.2.2	hessian-vector-product	39
6.3	CNNの物性計算への応用	40
6.3.1	CGCNN	41
6.3.2	R3DVS	42
第7章	3DCNNによる交換相関汎関数法の開発	45
7.1	3DCNNによる交換相関汎関数のモデリング	45
7.2	ネットワークの訓練	47
7.3	入力データ	50
第8章	実験	51
8.1	既存の汎関数のモデリング	51
8.2	既存汎関数でのバンドギャップの計算	53
8.3	小規模データセットでの検証	54
8.4	バンドギャップの補正	55
第9章	結論	58

目次	iv
謝辭	60
参考文献	61
对外发表	64

第1章

序論

1.1 研究背景

コンピュータ科学者のジェームズ・ニコラス・グレイは現在の人類の科学研究として「第四のパラダイム」を提唱した。¹ 第一のパラダイムでは観測したデータを経験的に処理していく経験記述的手法により科学研究が行われた。天動説や静電気、磁石の発見などがこれにあたる。第二のパラダイムでは観測データを数学や物理により解析し、背後に隠れている原理や法則を理解しようとする理論的手法による科学研究が行われた。ニュートンの運動方程式に始まり、電磁気学におけるマクスウェル方程式、熱力学におけるボルツマンの法則、アインシュタインによる相対性理論など現代科学の基礎理論がこのパラダイムにおいて生まれた。第三のパラダイムではコンピュータによるシミュレーションにより自然を数値計算的に理解しようとする計算処理的手法による科学研究が行われた。第二のパラダイムにおいて考案された方程式の中には解析的に解くことができないものがあつた。20世紀中盤に登場したコンピュータによってそれらに対して数値計算的に解を求めることが可能になり第三のパラダイムを象徴する計算科学が誕生した。現在、コンピュータの遍在性とその性能の飛躍的な向上により膨大な量の情報が世に溢れている。その情報を集約し、機械学習などの情報処理技術によって解析することで、新たな知見を得る情報集約型科学が第四のパラダイムとして提唱されている。

物性研究も第四のパラダイムを迎えており、長年にわたって蓄積されてきた材料データを集約し、機械学習などの情報処理技術により高いスループットで新規機能性材料を設計、開発する分野が誕生している。この分野は”マテリアルズインフォマティクス”と呼ばれ世界中で国を挙げて盛んに研究されている。

通常のマテリアルズインフォマティクスの研究では密度汎関数法などの第一原理計算手法を機械学習モデルによってバイパスすることを主眼に置いている。時間のかかる第一原理計算を行わずに機械学習モデルにより直接物性値を得ることができれば多くの物質の候補の中から有用なものを短時間で選び出すことができる。しかし、この過程で得られるのは有用な物質の候補だけであり、ブラックボックスである機械学習モデルから科学的知見が得られることは少ない。

本研究はこの通常のマテリアルズインフォマティクス研究とは異なるアプローチで物性データから新たな知見を生み出すことを目的とした。本研究の目標は物性データをもとに機械学習モデル3DCNNを訓練することで、第一原理計算手法である密度汎関数理論を補強する新たな交換相関汎関数を開発することである。交換相関汎関数は電子の多体量子相互作用のすべてが詰まっており、もし厳密な汎関数を得ることができれば物性科学のほぼすべてを知ることができるものである。この厳密な交換相関汎関数を機械学習モデルによって近似することができれば、そのモデルの入出力を調べることで電子相互作用に関する新たな科学的知見を生み出すことができる可能性がある。これはまさに第四のパラダイムにおける科学研究手法ということができる。

本研究により機械学習モデルから新たな知見を生み出すことができれば、自然の描像を追い求める科学研究において新たなアプローチを切り開くことになる。その意味でこの研究を行うことの意義は大きい。

1.2 本論文の構成

本論文ではまず2章において物性計算の舞台となる多体電子系に関してその基礎方程式とそれを解くための近似手法である断熱近似と一電子近似について記載する。また多体波動関数を一電子軌道によって記述するハートリー法、ハートリーフォック法を紹介しそれらに足りない効果、すなわち電子相関の概念について触れる。3章では本論文の主題である密度汎関数法についてその基となったトーマスフェルミ理論について述べた後、理論の基礎方程式であるホーエンベルグコーンの定理から基礎方程式であるコーンシャム方程式を導出する。4章ではコーンシャム方程式の導出過程で登場した交換相関汎関数について分類し、それぞれの特徴について述べる。5章では密度汎関数法の最大の弱点の一つであり本研究の改善すべき目標でもあるバンドギャップに関してその誤差の原因と改善法について見ていく。6章では本研究で用いる3DCNNの基礎となるニューラルネットワークに関して誤差逆伝播法を用いた各種演算と、畳み込みニューラルネットワークを物性研究に適応した研究について紹介する。7章では本研究で用いられる3DCNNモデルの概要と学習アルゴリズムについて記載し、8章ではそれを用いて実際に交換相関汎関数を改善する実験を行った結果と考察を述べる。9章では実験をまとめ、今後の展望を述べる。

第2章

物質中の電子

物質の性質を知る上で重要になる要素は電子である。電子は物質中において他の原子核や電子とクーロン力によって相互作用している。このクーロン相互作用によって原子同士の結合が生まれ、目に見える物質として存在していることになる。電子はその内部との相互作用だけでなく、外界からの刺激に対する反応においても大きな役割を果たす。物質に電場を印加した際に電気が流れるかどうかは、伝導電子の存在によって決まっている。また外部から物質に光を当てた際にどの波長の光が反射し、透過するかは物質中の電子と光の電場との結合具合によって決まる。他にも比熱や磁性、弾性率など多種多様な物性が電子によって支配されている。したがって、物性を知るためには物質中の電子の様相を知る必要がある。しかしながら物質中の電子は多くの原子核や電子と相互作用する多体量子系であり、このことが物性計算を行う上で最大の問題になる。三体以上の多体系は特別な場合を除いてその解析解を求めることができない。したがって多体系である電子の運動を知るためには、何らかの近似を行う必要がある。この章では多体量子系である物質中の電子の運動を実際に解くことができる形式に近似する手法について述べる。

2.1 断熱近似

物質は多数の原子核と電子によって構成されている。その性質は原子核、電子間に働くクーロン相互作用によって支配されているため、その相互作用の様相を知ることが物性解析の主題である。複数の原子核と電子によって構成される系のハミルトニアンは以下のように表される。

$$\begin{aligned} \mathcal{H} &= \mathcal{H}_e + \mathcal{H}_{en} + \mathcal{H}_n \\ &= -\sum_i \frac{\hbar^2}{2m} \nabla_i^2 + \frac{1}{2} \sum_{i \neq j} \frac{e^2}{|\mathbf{r}_i - \mathbf{r}_j|} - \sum_{i,n} \frac{Z_n e^2}{|\mathbf{r}_i - \mathbf{R}_n|} - \sum_n \frac{\hbar^2}{2M_n} \nabla_n^2 + \frac{1}{2} \sum_{n \neq m} \frac{Z_n Z_m e^2}{|\mathbf{R}_n - \mathbf{R}_m|} \end{aligned} \quad (2.1)$$

ここで \mathcal{H}_e は電子の運動エネルギーとクーロン相互作用、 \mathcal{H}_{en} は電子、原子核間のクーロン相互作用、 \mathcal{H}_n は原子核の運動エネルギーとクーロン相互作用を表す。この系のハミルトニアンに対する固有関数を Φ 、固有値を ε とすると、この系のシュレーディンガー方程式は

$$\mathcal{H}\Phi(\mathbf{r}_1, \dots, \mathbf{R}_1, \dots) = \varepsilon\Phi(\mathbf{r}_1, \dots, \mathbf{R}_1, \dots) \quad (2.2)$$

と表される。(2.2)を解くことができれば、原理的にはその物質に関する情報を完全に把握することができるが、この式を解析的、数値的に直接解くことは限られた系を除いて極めて困難な問題である。なぜならば、固有関数 Φ は原

子核,電子の個数のオーダーを持つ非常に多次元な関数であり,その多次元関数 Φ に二階の偏微分と多体相互作用を含むハミルトニアン \mathcal{H} が作用するため,この固有値問題は極めて多次元かつ複雑なものになるからである.

そこでこのシュレーディンガー方程式を解くことができるよう近似を行う. まず初めに原子核と電子の運動を分離する. 原子核を構成する陽子は電子の1840倍の質量を持つため,原子核の運動は電子の運動に比べ非常に遅く,それぞれを分離して考えても大きな問題にはならない. そこで固有関数 Φ を以下の二つの関数 Ψ と ϕ の積として表す.

$$\Phi(\mathbf{r}_1, \dots, \mathbf{R}_1, \dots) = \Psi(\mathbf{r}_1, \dots, \mathbf{R}_1, \dots)\phi(\mathbf{R}_1, \dots) \quad (2.3)$$

ここで Ψ は電子系のハミルトニアン \mathcal{H}_e の固有関数であり,

$$\mathcal{H}_e\Psi(\mathbf{r}_1, \dots, \mathbf{R}_1, \dots) = E(\mathbf{R}_1, \dots)\Psi(\mathbf{r}_1, \dots, \mathbf{R}_1, \dots) \quad (2.4)$$

を満たす. この固有関数にハミルトニアン(2.2)を作用させることにより,

$$\mathcal{H}\Phi = \Psi \left[-\sum_n \frac{\hbar^2}{2M_n} \nabla_n^2 + E(\mathbf{R}_1, \dots) + \frac{1}{2} \sum_{n \neq m} \frac{Z_n Z_m e^2}{|\mathbf{R}_n - \mathbf{R}_m|} \right] \phi - \sum_n \frac{\hbar^2}{2M_n} \{2\nabla_n \Psi \cdot \nabla_n \phi + \phi \nabla_n^2 \Psi\} \quad (2.5)$$

が得られる. (2.5)の右辺第二項は電子と原子の相互作用を表す項であり,この項を無視すると,

$$\left\{ -\sum_n \frac{\hbar^2}{2M_n} \nabla_n^2 + E(\mathbf{R}_1, \dots) + \frac{1}{2} \sum_{n \neq m} \frac{Z_n Z_m e^2}{|\mathbf{R}_n - \mathbf{R}_m|} \right\} \phi(\mathbf{R}_1, \dots) = \varepsilon \phi(\mathbf{R}_1, \dots) \quad (2.6)$$

となり,原子核系のシュレーディンガー方程式が導かれる. この近似により(2.2)を電子系のシュレーディンガー方程式(2.4)と原子核系のシュレーディンガー方程式(2.6)に分離することができる. この近似は電子系と原子核系の断熱的な分離を意味するので,断熱近似,またはボルン・オッペンハイマー近似と呼ばれる.

断熱近似によって得られた電子系のシュレーディンガー方程式を書き下すと,

$$\left[-\sum_i \frac{\hbar^2}{2m} \nabla_i^2 + \frac{1}{2} \sum_{i \neq j} \frac{e^2}{|\mathbf{r}_i - \mathbf{r}_j|} - \sum_{i,n} \frac{Z_n e^2}{|\mathbf{r}_i - \mathbf{R}_n|} \right] \Psi(\mathbf{r}_1, \dots, \mathbf{R}_1, \dots) = E(\mathbf{R}_1, \dots)\Psi(\mathbf{r}_1, \dots, \mathbf{R}_1, \dots) \quad (2.7)$$

となる. 物質の電気伝導や光学応答など物性分野で注目される現象の多くは電子のダイナミクスに起因するものであり,それを解明するためには(2.7)を解く必要がある. 断熱近似によって原子核の運動は分離することができたが,(2.7)には未だに電子同士のクーロン相互作用に起因する多体相互作用項が含まれており,直接解くことは容易ではない. この電子系のシュレーディンガー方程式を解くために今日まで様々な近似法が開発されてきた. 本論文の主題である密度汎関数法もこの電子の多体シュレーディンガー方程式を解くために開発されたものである. 次の章からはこの多体シュレーディンガー方程式を解くための近似法について述べる.

2.2 一電子近似

電子数が一つの系である水素原子のシュレーディンガー方程式は解析的に解くことができる. しかし電子数が二つ以上になるともはやシュレーディンガー方程式を解析的に解くことはできず,何かしらの近似に頼ることになる.

多体シュレーディンガー方程式(2.7)を解くことを困難にしている原因として,多体波動関数 $\Psi(\mathbf{r}_1, \dots, \mathbf{r}_N, \dots)$ の正確な形を知ることが難しいことがある. この困難を避けるための近似法として多体波動関数を一電子波動関数に分解する一電子近似がある. すなわち,多体波動関数 $\Psi(\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_N)$ を一電子波動関数 $\phi(\mathbf{r}_i)$ の集合により,

$$\Psi(\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_N) = f(\phi(\mathbf{r}_1), \phi(\mathbf{r}_2), \dots, \phi(\mathbf{r}_N)) \quad (2.8)$$

として表す. 一電子波動関数によって表された多体波動関数を多体シュレーディンガー方程式(2.7)に適用することで,

$$\begin{aligned} \mathcal{H}\phi_i(\mathbf{r}) &= \varepsilon_i\phi_i(\mathbf{r}) \\ \mathcal{H} &= \frac{\hbar^2}{2m}\nabla^2 + V_{\text{eff}}(\mathbf{r}) \end{aligned} \quad (2.9)$$

として一電子に関するシュレーディンガー方程式を得ることができる. 一電子シュレーディンガー方程式(2.9)における一電子ハミルトニアン第二項目は一つの電子が他の電子や原子核から平均的に感じるポテンシャルであり,平均場と呼ばれる. このことから一電子近似は平均場近似とも呼ばれる. もし多体波動関数を一電子軌道によって厳密に表すことができれば,多体シュレーディンガー方程式を一電子の問題として厳密に解くことができる. 今日まで様々な一電子近似法が開発されてきたが,未だに多体波動関数を一電子軌道で厳密に表すことは極めて限られた系を除いて実現的ではない. 精度の高い手法程,多くの一電子軌道の計算が必要になり,精度を求めるほど軌道の組み合わせの数が増えるため計算時間が爆発的に増大する. 次節から実際の科学計算でよく用いられるハートリー法とハートリーフォック法を紹介し,それらの手法に足りないもの,すなわち電子相関について述べる.

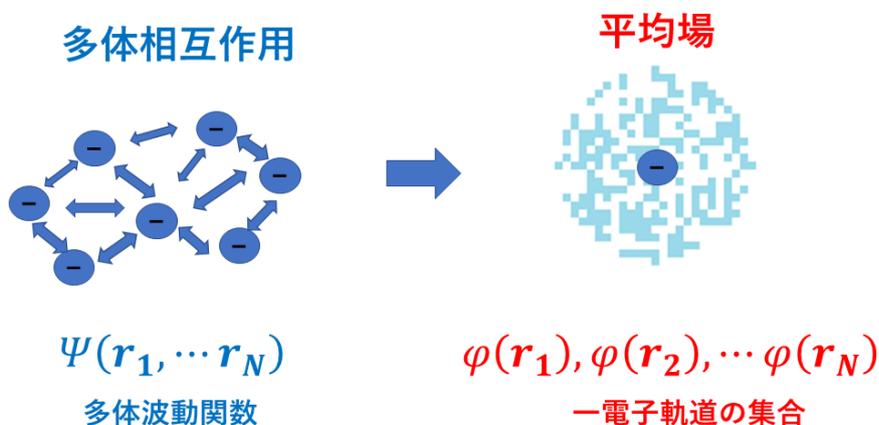


図2.1 一電子近似

2.2.1 ハートリー法

シュレーディンガー方程式が考案された翌々年の1928年,ハートリーは多体電子系に対するシュレーディンガー方程式の解法としてハートリー法を提案した. ハートリー法では N 電子系の多体波動関数 $\Psi(\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_N)$ を一電子軌道 $\phi_i(\mathbf{r}_i)$ の積として表す. すなわち

$$\Psi(\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_N) = \phi_1(\mathbf{r}_1)\phi_2(\mathbf{r}_2)\cdots\phi_N(\mathbf{r}_N) \quad (2.10)$$

ここで波動関数から原子核の座標 \mathbf{R}_i を省略した。原子核は電子の運動に比べ非常に遅いため、原子核座標は固定して考えても問題にはならない。以降では原子核からのポテンシャル項を単なる外部ポテンシャル $v_{\text{ext}}(\mathbf{r}_i)$ とし、原子核座標を省略して多体シュレーディンガー方程式(2.7)を書くことにする。多体シュレーディンガー方程式(2.7)のハミルトニアンは

$$\mathcal{H} = -\sum_i \frac{\hbar^2}{2m} \nabla_i^2 + \frac{1}{2} \sum_{i \neq j} \frac{e^2}{|\mathbf{r}_i - \mathbf{r}_j|} + \sum_i v_{\text{ext}}(\mathbf{r}_i) \quad (2.11)$$

$$v_{\text{ext}}(\mathbf{r}_i) = \sum_n \frac{Z_n e^2}{|\mathbf{r}_i - \mathbf{R}_n|}$$

となるが、このハミルトニアンに対しハートリー法における多体波動関数でエネルギーの期待値をとり、一電子軌道に関するエネルギーの変分が0であるとすると、

$$\left\{ -\sum_i \frac{\hbar^2}{2m} \nabla_i^2 + \sum_i v_{\text{ext}}(\mathbf{r}_i) + e^2 \sum_{j \neq i} \int \frac{|\phi_j(\mathbf{r}')|^2}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r}' \right\} \phi_i(\mathbf{r}) = \epsilon_i \phi_i(\mathbf{r}) \quad (2.12)$$

というハートリー法における一電子軌道の従うハートリー方程式が導かれる。ここでハートリー法によって導かれた平均場を見てみると、原子核のポテンシャル $v_{\text{ext}}(\mathbf{r}_i)$ と電子密度 $\rho(\mathbf{r})$ によるポテンシャル

$$V_c(\mathbf{r}) = e^2 \sum_{j \neq i} \int \frac{|\phi_j(\mathbf{r}')|^2}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r}' \quad (2.13)$$

からなっていることが分かる。この項は系の平均化された電子密度との古典的なクーロン相互作用であり、ハートリー項と呼ばれる。ハートリー法は物理的に妥当に思われる方程式を導くが、実際の系に対して計算を行うと全く使い物にならないことが分かった。これは一電子が感じる平均場が古典的なクーロン相互作用以外のポテンシャルを含んでいることが原因の一つである。次節で紹介するハートリーフォック法により導かれる平均場はハートリー法には存在しない量子統計に由来するポテンシャルが含まれる。

2.2.2 ハートリーフォック法

電子はスピン量子数1/2を持つフェルミオンである。複数のフェルミオンである電子によって構成される多体波動関数が満たすべき重要な性質として反対称性がある。すなわち多体波動関数の任意の電子対の座標変数の交換に対して多体波動関数はその符号を変えなければならない。ここでは位置座標とスピン座標をまとめて座標変数と呼ぶことにする。多体波動関数 Ψ に対して

$$\Psi(\cdots, \mathbf{r}_i, \cdots, \mathbf{r}_j, \cdots) = -\Psi(\cdots, \mathbf{r}_j, \cdots, \mathbf{r}_i, \cdots) \quad (2.14)$$

となるのが電子の反対称性から要請されるが、ここで $\mathbf{r}_i = \mathbf{r}_j$ とすると、 $\Psi = 0$ となることが分かる。これは多体電子系において電子が同時に同じ座標変数をとることができないということを表しており、パウリの排他原理として量子力学の基礎原理となっている。

ハートリー法では多体波動関数を一電子軌道の積によって表したがその場合多体波動関数は反対称性を満たさない。一方、ハートリーフォック法では電子の反対称性を満たすように多体波動関数を一電子軌道で近似することで、パウリの排他原理由来のポテンシャル項が導出される。ハートリーフォック法では多体波動関数 Ψ を一電子軌道 ϕ_i の行列式により、

$$\Psi(\mathbf{r}_1, \dots, \mathbf{r}_N) = \frac{1}{\sqrt{N!}} \begin{vmatrix} \phi_1(\mathbf{r}_1) & \cdots & \phi_1(\mathbf{r}_N) \\ \vdots & \ddots & \vdots \\ \phi_N(\mathbf{r}_1) & \cdots & \phi_N(\mathbf{r}_N) \end{vmatrix} \quad (2.15)$$

として表す。この多体波動関数はスレーター行列式と呼ばれ、電子の座標変数の交換に対して反対称性を満たす。多体電子系のハミルトニアン(2.7)に対してスレーター行列式によりエネルギー期待値を計算し、一電子軌道に対する変分を0とすると

$$\left\{ -\sum_i \frac{\hbar^2}{2m} \nabla_i^2 + \sum_i v_{\text{ext}}(\mathbf{r}_i) + e^2 \sum_{j=1}^N \int \frac{|\phi_j(\mathbf{r}')|^2}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r}' - e^2 \sum_{j=1(//)}^N \int \frac{\phi_j^*(\mathbf{r}')\phi_i(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r}' \right\} \phi_i(\mathbf{r}) = \epsilon_i \phi_i(\mathbf{r}) \quad (2.16)$$

という電子の反対称性を考慮した一電子近似方程式であるハートリーフォック方程式が導かれる。ハートリーフォック方程式における平均場を見てみるとハートリー法で導かれた平均場に加え、新たな項、

$$V_x(\mathbf{r}) = e^2 \sum_{j=1(//)}^N \int \frac{\phi_j^*(\mathbf{r}')\phi_i(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r}' \quad (2.17)$$

ここで(2.17)の和における(//)の意味は*i*番目の電子と平行なスピン座標についての和を表す。この項は電子の反対称性を考慮したことにより生じた効果であり、古典的な対応を持たない量子力学的な相互作用項である。この項は交換ポテンシャルと呼ばれ、この項によって表される電子間相互作用を交換相互作用と呼ぶ。交換相互作用は平行スピンをもつ電子同士がパウリの排他原理を満たすように、互いに接近することを妨げるように働く作用である。

交換相互作用は電子のスピンが大きな役割を果たす磁性を考える際は極めて重要な効果である。例として磁性を考える際によく使われるイジング模型を考える。イジング模型は二つのスピン量子数 $\sigma_i = \pm 1$ をもつ格子点からなり、それらの格子点の間に交換相互作用が働くモデルである。イジング模型におけるハミルトニアンは、

$$\mathcal{H} = \sum_{\langle i,j \rangle} J_{ij} \sigma_i \cdot \sigma_j \quad (2.18)$$

と表される。ここで(2.18)における J_{ij} は格子点*i*, *j*の間に働く交換相互作用の大きさを表す。もし $J_{i,j}$ 符号が負であれば*i*と*j*のスピンは平行、すなわち $\sigma_i \cdot \sigma_j = +1$ となった方がエネルギーが下がる。逆に $J_{i,j}$ が正であった場合、*i*と*j*のスピンは反平行、すなわち $\sigma_i \cdot \sigma_j = -1$ となった方がエネルギーが下がる。すなわち全体として*J*が負であれば強磁性的であり、正であれば反強磁性的になることが分かる。このように物質の磁性は純粋な量子力学的効果である交換相互作用がマクロなスケールに表れた巨視的量子現象である。

ここで再びハートリーフォック方程式(2.16)における平均場を考える。ハートリー方程式(2.12)では電子の古典的なクーロン相互作用であるハートリー項に関する和は*j* = *i*の項は除外していたが、ハートリーフォック方程式で

はすべての軌道に関する和になっている。これは $j = i$ の項はハートリー項と交換相互作用項で打ち消しあうため、和に含めても問題ないためである。このクーロン相互作用と交換相互作用の打ち消しあいは、後に議論する密度汎関数法における自己相互作用に深く関わっている。

ハートリーフォック方程式は一電子近似にパウリの排他原理を導入することに成功し、量子力学の基礎理論となったが、実際の物質に適用すると、十分な化学的精度を得ることができないことが分かった。この原因はハートリーフォック方程式が見積もることのできなかつた多体相互作用効果に由来するものあり、次節ではこの効果について述べる。

2.3 電子相関

ハートリーフォック方程式を実際の物質の計算に適用すると現実にはありえない結果になることがある。例えば電子密度が希薄な系では電子のスピン向きが揃いやすくなり、実際は非磁性の物質が計算上では強磁性体になってしまう。またフェルミ準位での状態密度が有限ではなく0になるという重大な欠陥を抱えており、ハートリーフォック法を固体の計算に適用するとバンドギャップを過大評価してしまい使い物にならない。² この問題はハートリーフォック法では見積もることのできない残りたった1%未満のエネルギー誤差に起因している。

1955年、レフディンはこのエネルギー誤差を電子相関と定義した。³ ハートリーフォック法によって見積もられるエネルギーを E_{HF} 、厳密な物質のエネルギーを E_{exact} とすると、電子相関によるエネルギーは

$$E_{\text{corr}} = E_{\text{HF}} - E_{\text{exact}} \quad (2.19)$$

と定義される。電子相関が生じる原因として、ハートリーフォック法では反平行スピンを持つ電子間の相互作用が見積もられていないことが挙げられる。多体シュレーディンガー方程式(2.7)におけるハミルトニアンは二つの電子の距離 $r_{ij} = |\mathbf{r}_i - \mathbf{r}_j|$ の関数であり $r_{ij} = 0$ で特異点を持つ。しかしながら系のエネルギー E は有限であるため、多体波動関数 $\Psi(\mathbf{r}_1, \dots)$ は $r_{ij} = 0$ で0とならなければならない。すなわち多体波動関数 $\Psi(\mathbf{r}_1, \dots)$ は $r_{ij} = 0$ で孔を持っていると考えられ、この孔をクーロン孔と呼ぶ。⁴ ハートリーフォック法では平行スピンを持つ電子間には交換相互作用により互いに接近しないよう退け合う力が働く。すなわち電子は自身の周りに同じ向きを持つ電子が入れない領域を伴って運動しており、この領域のことをフェルミ孔と呼ぶ。⁴ これは二つの電子間に働く直接的なクーロン相互作用とは異なり、パウリの排他原理に由来する量子統計的な効果である。一方反平行スピンを持つ電子間の運動は互いに独立で相関していない。そのため反平行スピンを持つ電子間にはクーロン孔が存在せず、上記の議論と矛盾する。この反平行スピンを持つ電子同士のクーロン孔の欠如が電子相関の最大の原因である。⁵

2.3.1 電子相関を取り込むための手法

一電子近似に電子相関を取り込むための指針として電子配置間相互作用を考慮することが挙げられる. このことを考える例として水素分子の基底状態の結合解離を考える.

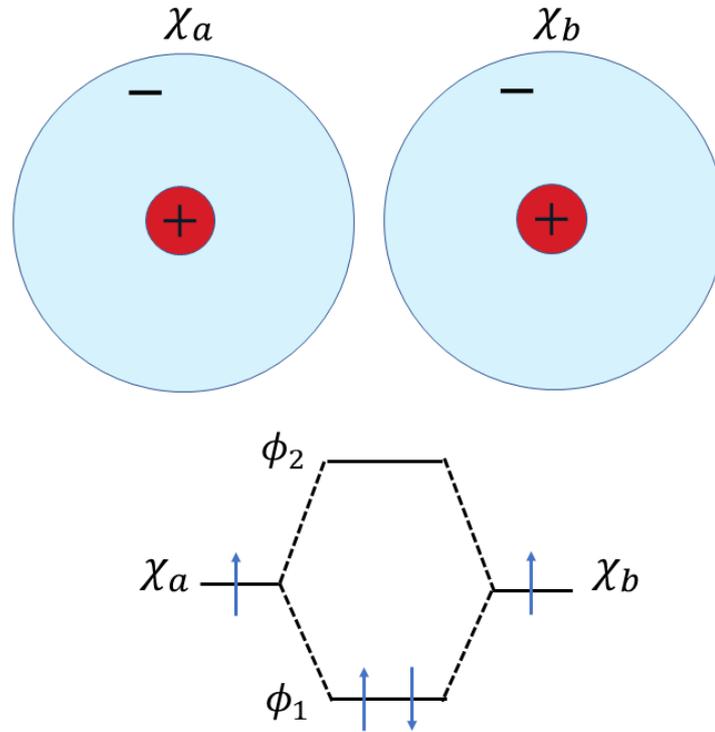


図2.2 水素原子

水素原子の原子軌道を χ_A, χ_B とすると,この二つの原子軌道から結合性分子軌道

$$\phi_1 = \frac{1}{\sqrt{2 + 2S_{AB}}}(\chi_A + \chi_B) \quad (2.20)$$

と反結合性分子軌道

$$\phi_2 = \frac{1}{\sqrt{2 - 2S_{AB}}}(\chi_A - \chi_B) \quad (2.21)$$

が作られる.ここで S_{AB} は二つの軌道の重なり積分であり,

$$S_{AB} = \int d^3\mathbf{r} \chi_A^*(\mathbf{r})\chi_B(\mathbf{r}) \quad (2.22)$$

と計算される. ハートリーフォック法では,水素分子の二つの電子はエネルギーの低い結合性軌道に入っているの
で,電子の多体波動関数はスレーター行列式により,

$$\Psi_{\text{gnd}}(\mathbf{r}_1, \mathbf{r}_2) = \frac{1}{\sqrt{2}}\det(\phi_1(\mathbf{r}_1)\phi_1(\mathbf{r}_2)) \quad (2.23)$$

となる. この多体波動関数は $\text{H}_2 \rightarrow 2\text{H}$ の解離極限では重なり積分が $S_{AB} = 0$ となるので

$$\Psi_{\text{gnd}}(\mathbf{r}_1, \mathbf{r}_2) = \frac{1}{2\sqrt{2}}\det(\chi_A(\mathbf{r}_1)\chi_B(\mathbf{r}_2) + \chi_B(\mathbf{r}_1)\chi_A(\mathbf{r}_2) + \chi_A(\mathbf{r}_1)\chi_A(\mathbf{r}_2) + \chi_B(\mathbf{r}_1)\chi_B(\mathbf{r}_2)) \quad (2.24)$$

となる。しかしこの多体波動関数は現実の解離極限を正しく表していない。解離極限における多体波動関数の前半の二項はそれぞれの原子軌道に一つずつ電子が収まった共有結合性状態であるが、後半の二項はどちらかの原子軌道に二つの電子が収まった状態であり、極めて不安定な電子状態である。この問題は励起電子配置まで考慮することで解消することができる。反結合性軌道 ϕ_2 により表された励起電子配置のスレーター行列式は

$$\Psi_{\text{exc}}(\mathbf{r}_1, \mathbf{r}_2) = \frac{1}{\sqrt{2}} \det(\phi_2(\mathbf{r}_1)\phi_2(\mathbf{r}_2)) \quad (2.25)$$

となるがこの電子配置は解離極限において、

$$\Psi_{\text{exc}}(\mathbf{r}_1, \mathbf{r}_2) = \frac{1}{2\sqrt{2}} \det(\chi_A(\mathbf{r}_1)\chi_B(\mathbf{r}_2) + \chi_B(\mathbf{r}_1)\chi_A(\mathbf{r}_2) - \chi_A(\mathbf{r}_1)\chi_A(\mathbf{r}_2) - \chi_B(\mathbf{r}_1)\chi_B(\mathbf{r}_2)) \quad (2.26)$$

という形になる。前半の二項は基底電子配置(2.23)のものと同じであるが、後半の二項は基底電子配置のものとは符号が反転している。したがって、基底電子配置と励起電子配置の和をとることで基底状態の多体波動関数を表すと

$$\begin{aligned} \Psi(\mathbf{r}_1, \mathbf{r}_2) &= \frac{1}{2\sqrt{2}} \det(\phi_1(\mathbf{r}_1)\phi_1(\mathbf{r}_2) - \phi_2(\mathbf{r}_1)\phi_2(\mathbf{r}_2)) \\ &= \frac{1}{2\sqrt{2}} \det(\chi_A(\mathbf{r}_1)\chi_B(\mathbf{r}_2) + \chi_B(\mathbf{r}_1)\chi_A(\mathbf{r}_2)) \end{aligned} \quad (2.27)$$

となり非物理的な項が消滅し、不安定化が解消される。

一般に電子状態はスレーター行列式で与えられる電子配置の線形和として

$$\Psi(\mathbf{r}_1, \dots) = \sum_I C_I \Phi_I(\mathbf{r}_1, \dots) \quad (2.28)$$

として書ける。空間、スピン変数の同じ電子配置のエネルギー準位が近づくと電子配置は混合し、エネルギーの安定化が起こる。この安定化したエネルギーを静的電子相関という。

このように複数の電子配置を考慮することにより電子相関を取り込むことができる。基底電子配置と励起電子配置を線形結合することで電子相関を取り込む方法を配置間相互作用(CI)法という。すべての電子配置を考慮する完全CI(full-CI)法を計算できれば、電子相関を完全に取り込むことができるため、理論的には厳密な計算を行うことができる。しかし、考え得る一電子軌道は有限であり、また電子数が増えるほど、考慮すべき電子配置は指数関数的に増大するため、full-CI計算を行うことは、限られた系を除いて現実的ではない。CI法以外にも結合クラスター(CC)法、メラプリセット(MP)摂動法⁶など今まで様々な手法が開発されてきた。多体シュレーディンガー方程式の真の解に近づくためには、複数の電子配置を効率的に取り込むことができる計算手法が必要になる。いかにして実効的な時間内で高度に電子相関を取り込むかが量子化学計算における重要な課題である。

第3章

密度汎関数法

先の章では,多体波動関数を一電子軌道により近似する手法について述べた. 多体波動関数を近似するうえで問題となるのが,真の多体波動関数に近づくためには,多くの一電子軌道に対する計算が必要になり,高度に電子相関を取り込む理論になるほど計算時間が増大することであった. この章で述べる密度汎関数法は多体波動関数の近似を経由することなく多体シュレーディンガー方程式を厳密に一電子軌道に書き換える手法である. ,密度汎関数法はその精度の高さと計算コストの少なさから,物性物理学,量子化学をはじめ生命科学,⁷地学,⁸天文学⁹など様々な分野で用いられ,大きな成功を収めている.¹⁰ この章ではまず密度汎関数法が誕生するきっかけとなったトーマスフェルミ理論について述べ,次に密度汎関数法の基礎定理であるホーエンベルグコーンの定理から基礎方程式であるコーンシャム方程式を導出する.. 最後に密度汎関数法を実際に計算するための手法について述べる.

3.1 トーマスフェルミ理論

シュレーディンガー方程式が提案された翌年の1927年,トーマスは固体結晶中の多電子系のシュレーディンガー方程式の近似解法として,電子密度に基づくエネルギー汎関数を開発した. この近似の重要な点は取り扱いが困難な多体波動関数の近似を行わず,電子密度により多電子系の情報を記述しようとしたことであり,この試みは密度汎関数法の原型になった. トーマスは電子密度によるシュレーディンガー方程式の解を導出するために固体中の電子を均一な電子ガスとして近似し,この近似の下で電子の運動エネルギーを電子密度の汎関数として

$$\begin{aligned} T^{TF} &= C_F \int d^3\mathbf{r} \rho^{\frac{5}{3}}(\mathbf{r}) \\ C_F &= \frac{3}{10} (3\pi^2)^{\frac{2}{3}} \end{aligned} \quad (3.1)$$

と表した. 翌年,フェルミはトーマスとは独立に絶対零度におけるフェルミ統計をもとにこの運動エネルギー汎関数を提案し,トーマスフェルミ理論を完成させた. しかしながらトーマスフェルミ理論は現実の物質に対して定量的な議論はおろか,定性的な議論すらできない. ディラックはトーマスフェルミ理論の不完全さは先の章で登場した交換相互作用(2.17)が考慮されていないことが原因であるとし,交換相互作用を電子密度の汎関数として

$$E_x = -\frac{3}{4} \left(\frac{3}{\pi} \right)^{\frac{1}{3}} \int d^3\mathbf{r} \rho^{\frac{4}{3}}(\mathbf{r}) \quad (3.2)$$

と表した交換汎関数を提案した。さらにワイツゼッカーは電子密度の勾配を用いた補正項を加えることで理論の改善を試みた。

$$T^W = \frac{1}{8} \int d^3\mathbf{r} \frac{|\nabla\rho(\mathbf{r})|^2}{\rho(\mathbf{r})} \quad (3.3)$$

のちにこの値の1/9がトーマスフェルミ理論の厳密な補正項であることが証明された。トーマスフェルミ理論はシュレーディンガー方程式を多体波動関数の近似を経由せず解く手法として画期的であったが、十分な精度を達成することができず、また物理的な裏付けもなかったため密度汎関数法が登場するまで忘れ去られることとなった。

3.2 ホーエンベルグコーンの定理

トーマスフェルミ理論が提案されてから40年弱経った1964年、トーマスフェルミ理論を物理的に裏付ける理論がホーエンベルグとコーンによって提案された。¹¹ ホーエンベルグコーンの定理の登場によって電子密度の情報から系の物理量を記述する密度汎関数法が誕生した。ホーエンベルグコーンの定理は二つの定理からなる。

3.2.1 ホーエンベルグコーンの第一定理

ホーエンベルグコーンの定理をなす一つ目の定理は

「ある系の基底状態の電子密度 ρ が決まると、それを基底状態に持つ外部ポテンシャルが定数倍の差を除いてただ一通りに求まる。」

というものである。電子の従う多体シュレーディンガー方程式を振り返る。

$$\begin{aligned} \mathcal{H} &= T + U + V \\ T &= - \sum_i \frac{\hbar^2}{2m} \nabla_i^2 \\ U &= \frac{1}{2} \sum_{i \neq j} \frac{e^2}{|\mathbf{r}_i - \mathbf{r}_j|} \\ V &= \sum_i v_{\text{ext}}(\mathbf{r}_i) = \sum_i \sum_n \frac{Z_n e^2}{|\mathbf{r}_i - \mathbf{R}_n|} \end{aligned} \quad (3.4)$$

ここで電子の運動エネルギー、電子間クーロン相互作用、原子核からのポテンシャル（外部ポテンシャル）をそれぞれ T, U, V とした。系特有の情報は外部ポテンシャル V のみであり、外部ポテンシャル V からシュレーディンガー方程式の解である多体波動関数やエネルギー固有値が導かれる。電子密度 $\rho(\mathbf{r})$ は多体波動関数 $\Psi(\mathbf{r}_1, \dots)$ から、

$$\rho(\mathbf{r}) = N \int \Psi^*(\mathbf{r}, \mathbf{r}_2, \dots, \mathbf{r}_N) \Psi(\mathbf{r}, \mathbf{r}_2, \dots, \mathbf{r}_N) d\mathbf{r}_2 \cdots \mathbf{r}_N \quad (3.5)$$

と導かれるため電子密度 ρ は外部ポテンシャル V と

$$V \mapsto \rho \quad (3.6)$$

という関係になる。ホーエンベルグコーンの第一定理では系の基底状態における電子密度 ρ についてこの逆写像、すなわち

$$\rho \mapsto V \quad (3.7)$$

が成り立つことを保証する. この定理を証明する.

まず一つの波動関数 Ψ が定数部分以外で異なる外部ポテンシャル V, V' から作られる二つのハミルトニアン

$$\begin{aligned}\mathcal{H} &= \mathcal{H}_0 + V \\ \mathcal{H}' &= \mathcal{H}_0 + V'\end{aligned}\quad (3.8)$$

の両方の固有状態にならないこと背理法によって示す. Ψ が両方のハミルトニアンの固有状態であるとする,

$$\begin{aligned}\mathcal{H}\Psi &= E\Psi \\ \mathcal{H}'\Psi &= E'\Psi\end{aligned}\quad (3.9)$$

となる. ここで E, E' はそれぞれの固有エネルギーである. 上式(3.9)の差分をとると,

$$(V - V')\Psi = (E - E')\Psi \quad (3.10)$$

となるが,波動関数 Ψ は恒等的に0になることはないため,

$$V - V' = E - E' \quad (3.11)$$

となる. 固有エネルギー E, E' が定数であることから, $V - V'$ も定数となるが,これは外部ポテンシャル V, V' が定数部分以外で異なるということに矛盾する. したがって波動関数 Ψ は二つのハミルトニアン $\mathcal{H}, \mathcal{H}'$ の固有関数に同時になることはなく,波動関数とハミルトニアンは一対一対応する. ハミルトニアンの系に固有な部分は外部ポテンシャル V のみであるから,このことは Ψ と外部ポテンシャル V が一対一対応することを示している. すなわち, Ψ が与えられれば,その Ψ を生じる外部ポテンシャル V は定数倍の差を除いて一意に定まる. このことを用いて電子密度 ρ と外部ポテンシャル V が一対一対応することを示す. 先の議論におけるハミルトニアン \mathcal{H} の基底状態の波動関数を Ψ とする. $\mathcal{H} = \mathcal{H}' + V - V'$ であるから,基底状態のエネルギー E は

$$\begin{aligned}E &= \langle \Psi | \mathcal{H} | \Psi \rangle \\ &= \langle \Psi | \mathcal{H}' | \Psi \rangle + \langle \Psi | (V - V') | \Psi \rangle \\ &= \langle \Psi | \mathcal{H}' | \Psi \rangle + \int \{V(\mathbf{r}) - V'(\mathbf{r})\} \rho(\mathbf{r}) dV\end{aligned}\quad (3.12)$$

となる. 基底状態の波動関数 Ψ はこの $E = \langle \Psi | \mathcal{H} | \Psi \rangle$ を最小化するものである. ハミルトニアン \mathcal{H}' の基底状態の波動関数を Ψ' とし, Ψ と Ψ' が同一の電子密度 ρ を与えると仮定する. このときハミルトニアン \mathcal{H} の波動関数 Ψ' に対するエネルギー期待値は

$$\begin{aligned}\langle \Psi' | \mathcal{H} | \Psi' \rangle &= \langle \Psi' | \mathcal{H}' | \Psi' \rangle + \langle \Psi' | (V - V') | \Psi' \rangle \\ &= \langle \Psi' | \mathcal{H}' | \Psi' \rangle + \int \{V(\mathbf{r}) - V'(\mathbf{r})\} \rho(\mathbf{r}) dV\end{aligned}\quad (3.13)$$

となる. (3.12)と(3.13)を比較すると, Ψ' が \mathcal{H}' の基底状態のハミルトニアンであることから

$$\langle \Psi | \mathcal{H} | \Psi \rangle \geq \langle \Psi' | \mathcal{H} | \Psi' \rangle \quad (3.14)$$

となる. Ψ はハミルトニアン \mathcal{H} の基底状態であることからこの等号が成り立たねばならない. したがって, $\Psi = \Psi'$ となるか, Ψ と Ψ' はハミルトニアン \mathcal{H}' に対する縮退した基底状態ということになる. 先の議論より波動関数 Ψ と外

部ポテンシャル V は一対一対応するため,縮退のあるなしにかかわらず基底状態の電子密度 ρ に対しての外部ポテンシャル V は一意に定まることになる. これにより,ホーエンベルグコーンの第一定理(3.7)が証明された.

この定理の重要な主張は電子密度が系の情報を完全に含んでいるということである. 電子密度が与えられれば,それを基底状態とする外部ポテンシャルが求まり,それをもとにシュレーディンガー方程式を解けば波動関数やエネルギー固有値を得ることができる. 波動関数が分かれば系のあらゆる物理量を導くことができるので,すべての物理量は電子密度の汎関数であるといえる. この定理により系のエネルギーを電子密度の汎関数として記述しようとしたトーマスフェルミ理論の正当性が証明され,密度汎関数法が物理的な裏付けを持って誕生することとなった.

3.2.2 ホーエンベルグコーンの第二定理

ホーエンベルグコーンの二つ目の定理は

「系のエネルギー固有値を最小にする電子密度はその系のシュレーディンガー方程式の真の解の電子密度となる」というものである. この定理は電子密度に対する変分原理として考えることができる. この定理を証明するには波動関数に対する変分原理であるシュレーディンガーリッツの変分原理を用いる. シュレーディンガーリッツの変分原理は N 電子系の規格化条件

$$\langle \Psi | \Psi \rangle = N \quad (3.15)$$

を満たす任意の波動関数 Ψ は系のエネルギー汎関数 $E_0(V, [\Psi])$ を

$$E_0(V, [\Psi]) = \langle \Psi | \mathcal{H} | \Psi \rangle \quad (3.16)$$

と定義すると,それは真の基底状態のエネルギー E_0 に比べて

$$E_0(V, [\Psi]) \geq E_0 \quad (3.17)$$

という関係が常に成り立ち,そしてこの等号が成立するのは Ψ が真の基底状態 Ψ_0 であるときに限られる,という原理である. エネルギー汎関数 $E_0(V, [\Psi])$ は外部ポテンシャル V という状況下で Ψ を試行関数とする汎関数であるという意味である. シュレーディンガー方程式の真の解 Ψ_0 はエネルギー汎関数の停留条件として求められるが,この条件はシュレーディンガー方程式そのものになる. シュレーディンガーリッツの変分原理は波動関数に対するものであったが,これを電子密度に対して書き換えたものがホーエンベルグコーンの第二定理となる密度変分原理である. この原理はシュレーディンガーリッツの変分原理とホーエンベルグコーンの第一定理により導かれる.¹² ホーエンベルグコーンの第一定理より,任意の電子密度 ρ に対し,その電子密度に対応する外部ポテンシャル V が一意に決まる. その外部ポテンシャルより導かれる波動関数を Ψ とし,系のシュレーディンガー方程式の外部ポテンシャルを V_0 ,その真の解に対する波動関数を Ψ_0 とする. $V = V_0$ である場合は,外部ポテンシャルと電子密度が一対一対応することから, ρ はシュレーディンガー方程式の真の解の電子密度となる. $V \neq V_0$ である場合は, Ψ は真の解である Ψ_0 とは決して同じものになることはなく,シュレーディンガーリッツの変分原理より,

$$E_0(V, [\Psi]) > E_0(V, [\Psi_0]) = E_0 \quad (3.18)$$

となるが、波動関数 Ψ は電子密度 ρ から導かれるため $E_0(V, [\Psi]) = E_0(V, [\rho])$ として考えられる。これを $V = V_0$ の場合と合わせると

$$E_0(V, [\rho]) \geq E_0 \quad (3.19)$$

という電子密度に対する変分原理が得られる。この密度変分原理により電子密度の真の解を探索するためには、系のエネルギーの電子密度に対する停留条件を調べればよいことが分かる。この停留条件から密度汎関数法の基礎方程式となるコーンシャム方程式が導かれることになる。¹³

3.3 コーンシャム方程式

ホーエンベルグコーンの第一定理により系のあらゆる物理量は電子密度の汎関数であることが示された。したがって系のエネルギー $E = \langle \Psi | \mathcal{H} | \Psi \rangle$ も電子密度の汎関数であるため、

$$\begin{aligned} E[\rho(\mathbf{r})] &= \langle \Psi | \mathcal{H} | \Psi \rangle \\ &= \langle \Psi | (T + U + V) | \Psi \rangle \\ &= F[\rho(\mathbf{r})] + \int V(\mathbf{r})\rho(\mathbf{r})d\mathbf{r} \\ F[\rho(\mathbf{r})] &= \langle \Psi | (T + U) | \Psi \rangle \end{aligned} \quad (3.20)$$

と表すことができる。¹⁴ ここで、エネルギー $E[\rho(\mathbf{r})]$ を外部ポテンシャルを含み系に特有である部分 $\int V(\mathbf{r})\rho(\mathbf{r})d\mathbf{r}$ と、系によらない部分 $F[\rho(\mathbf{r})]$ に分けた。この系によらない部分を電子密度の汎関数として書き下すことができれば、ホーエンベルグコーンの第二定理より、エネルギーの電子密度に関する停留条件を求めることで電子密度の真の解を得ることができる。しかしこの $F[\rho(\mathbf{r})]$ の電子密度 ρ に対する依存性を陽に書き表す方法は現在でも知られていない。この項を変分原理で取り扱える形にするため以下のように書き下す。

$$F[\rho(\mathbf{r})] = T_s[\rho(\mathbf{r})] + \frac{e^2}{2} \int \frac{\rho(\mathbf{r})\rho(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r}d\mathbf{r}' + E_{xc}[\rho(\mathbf{r})] \quad (3.21)$$

ここで右辺第一項 $T_s[\rho(\mathbf{r})]$ は相互作用のない系における電子の運動エネルギーであり、第二項は古典的なクーロン相互作用項である。そして第三項は前半の二項で含めることのできなかつた電子の相互作用に起因する運動エネルギーとクーロン相互作用の補正項であり、電子の交換相互作用と相関相互作用を含んでいる。この項によって記述されるエネルギーを交換相関エネルギーという。交換相関エネルギーに電子の多体効果の問題をすべて押し込んだため、この書き換えは近似を含まず厳密なもののみである。交換相関エネルギーは電子間相互作用によって生じる物理量であり、電子密度の汎関数としてその存在は保証されている。

この書き下された系のエネルギー汎関数(3.21)に対して電子密度 ρ による変分を施す。 N 電子系の電子密度 $\rho(\mathbf{r})$ は電子数に対する制約条件

$$N = \int \rho(\mathbf{r})d\mathbf{r} \quad (3.22)$$

を満たさねばならないので、ラグランジュの未定乗数を μ とすると系のエネルギー E の電子密度 $\rho(\mathbf{r})$ に対する停留条

件は

$$\frac{\delta}{\delta\rho(\mathbf{r})} \left[E[\rho(\mathbf{r})] - \mu \left(\int \rho(\mathbf{r}) - N \right) \right] = 0 \quad (3.23)$$

となる.ここで(3.21)を用いると,(3.23)は

$$\frac{\delta T_s[\rho]}{\delta\rho(\mathbf{r})} + V(\mathbf{r}) + e^2 \int \frac{\rho(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r}' + \frac{\delta E_{xc}}{\delta\rho(\mathbf{r})} - \mu = 0 \quad (3.24)$$

と展開することができる. 相互作用のない系における運動エネルギー T_s の電子密度 $\rho(\mathbf{r})$ に対する汎関数微分を求めるために,以下のような平均場 V_{eff} を持つ一電子方程式を考える.

$$\begin{aligned} \left[-\frac{\hbar^2}{2m} \nabla^2 + V_{\text{eff}}(\mathbf{r}) \right] \phi_i(\mathbf{r}) &= \varepsilon_i \phi_i(\mathbf{r}) \\ \rho(\mathbf{r}) &= \sum_1^N |\phi_i(\mathbf{r})|^2 \end{aligned} \quad (3.25)$$

この一電子方程式(3.25)が解ければ, T_s は一電子軌道 ϕ_i により

$$\begin{aligned} T_s[\rho(\mathbf{r})] &= \sum_{i=1}^N \int \phi_i^*(\mathbf{r}) \left(-\frac{\hbar^2}{2m} \nabla^2 \right) \phi_i(\mathbf{r}) d\mathbf{r} \\ &= \sum_{i=1}^N \varepsilon_i - \int V_{\text{eff}}(\mathbf{r}) \rho(\mathbf{r}) d\mathbf{r} \end{aligned} \quad (3.26)$$

と表すことができる.電子密度 $\rho(\mathbf{r})$ に対する変分を調べるために,

$$\begin{aligned} \rho(\mathbf{r}) &\rightarrow \rho(\mathbf{r}) + \delta\rho(\mathbf{r}) \\ V_{\text{eff}}(\mathbf{r}) &\rightarrow V_{\text{eff}}(\mathbf{r}) + \delta V_{\text{eff}}(\mathbf{r}) \\ \phi(\mathbf{r}) &\rightarrow \phi(\mathbf{r}) + \delta\phi(\mathbf{r}) \\ \varepsilon &\rightarrow \varepsilon + \delta\varepsilon \end{aligned} \quad (3.27)$$

という変化を考え,(3.25)に代入し,二次以上の変化分を無視すると,

$$\left[-\frac{\hbar^2}{2m} \nabla^2 + V_{\text{eff}}(\mathbf{r}) \right] \delta\phi_i(\mathbf{r}) + \delta V_{\text{eff}}(\mathbf{r}) \phi_i(\mathbf{r}) = \varepsilon_i \delta\phi_i(\mathbf{r}) + \delta\varepsilon_i \phi_i(\mathbf{r}) \quad (3.28)$$

とできて,(3.28)の左から $\phi_i(\mathbf{r})$ を掛けて積分すると

$$\varepsilon_i \int \phi_i^*(\mathbf{r}) \delta\phi_i(\mathbf{r}) d\mathbf{r} + \int \delta V_{\text{eff}}(\mathbf{r}) |\phi_i(\mathbf{r})|^2 = \varepsilon_i \int \phi_i^*(\mathbf{r}) \delta\phi_i(\mathbf{r}) d\mathbf{r} + \delta\varepsilon_i \int |\phi_i(\mathbf{r})|^2 \quad (3.29)$$

となる.さらに i を1から N まで和をとると

$$\sum_{i=1}^N \delta\varepsilon_i = \int \delta V_{\text{eff}}(\mathbf{r}) \rho(\mathbf{r}) d\mathbf{r} \quad (3.30)$$

という式が得られる.(3.26)を電子密度 $\rho(\mathbf{r})$ に関して汎関数微分し,この結果を使うと

$$\begin{aligned} \frac{\delta T_s[\rho]}{\delta\rho(\mathbf{r})} &= \sum_{i=1}^N \frac{\delta\varepsilon_i}{\delta\rho(\mathbf{r})} - \int \frac{\delta V_{\text{eff}}}{\rho(\mathbf{r})} \rho(\mathbf{r}) d\mathbf{r} - V_{\text{eff}}(\mathbf{r}) \\ &= -V_{\text{eff}}(\mathbf{r}) \end{aligned} \quad (3.31)$$

が得られる.この運動エネルギーの汎関数微分(3.31)を停留条件式(3.24)に代入することで,平均場 V_{eff} として

$$V_{\text{eff}}(\mathbf{r}) = V(\mathbf{r}) + e^2 \int \frac{\rho(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r}' + \frac{\delta E_{xc}[\rho]}{\delta\rho(\mathbf{r})} - \mu \quad (3.32)$$

が導かれる。こうして平均場 V_{eff} が求められたことにより、多体シュレーディンガー方程式を一電子問題として解くことができるようになった。まとめると

$$\begin{aligned} \left[-\frac{\hbar^2}{2m} \nabla^2 + V_{\text{eff}}(\mathbf{r}) \right] \phi_i(\mathbf{r}) &= \varepsilon_i \phi_i(\mathbf{r}) \\ V_{\text{eff}}(\mathbf{r}) &= V(\mathbf{r}) + e^2 \int \frac{\rho(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r}' + \frac{\delta E_{\text{xc}}[\rho]}{\delta \rho(\mathbf{r})} \\ \rho(\mathbf{r}) &= \sum_1^N |\phi_i(\mathbf{r})|^2 \end{aligned} \quad (3.33)$$

となる。この一電子方程式(3.33)をコーンシャム方程式といい、密度汎関数法を実際の計算として実行可能な形式に書き換えたものになるものになる。コーンシャム方程式で得られた平均場(3.32)に含まれる

$$V_{\text{xc}} = \frac{\delta E_{\text{xc}}[\rho]}{\delta \rho(\mathbf{r})} \quad (3.34)$$

の項は交換相関ポテンシャルと呼ばれ、交換相関エネルギーを与える電子の多体効果に由来するポテンシャルである。コーンシャム方程式を解き電子密度 ρ を手に入れることができれば、系のエネルギー E は

$$E[\rho] = \sum_{i=1}^N \delta \varepsilon_i - \frac{e^2}{2} \int \frac{\rho(\mathbf{r})\rho(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r}d\mathbf{r}' + E_{\text{xc}}[\rho(\mathbf{r})] - \int \frac{\delta E_{\text{xc}}}{\delta \rho(\mathbf{r})} \rho(\mathbf{r}) d\mathbf{r} \quad (3.35)$$

と求められる。ここでラグランジュの未定乗数 μ を無視しているが、無限結晶系の場合は真空準位が定まらず平均場の絶対的な原点に任意性がある。そのため μ を0としても問題はない。

密度汎関数法は多体シュレーディンガー方程式を厳密な形で一電子問題に書き換えることに成功したが、その基礎方程式であるコーンシャム方程式には交換相関汎関数という具体的な表式が分からない項が含まれたままである。この交換相関汎関数は未だ厳密な表式が分かっておらず、結局は何らかの手法で近似を行わなければならない。しかし交換相関汎関数の表式を考えることは、多体波動関数の様相を知ることに比べ格段に易しい問題である。なぜならば多体波動関数は電子の個数のオーダーを持つ非常に高次元な関数でありその実態を捉えることは困難であるが、交換相関汎関数はどのような系であっても三次元の電子密度の汎関数であるので取り扱いは多体波動関数に比べはるかに容易である。密度汎関数法の最大の功績の一つは多体電子系の解法として多体波動関数の近似とは別のアプローチを提示したことである。交換相関汎関数には古典的には捉えられなかった多体電子相互作用のすべてが詰まっており、その意味でこの汎関数は物性物理学の神髄ともいうことができる存在である。

3.4 コーンシャム方程式の解法

先の節で密度汎関数法の基礎方程式であるコーンシャム方程式を導出したが、この方程式をどのようにして解くかは計算科学において重要な問題である。この節ではコーンシャム方程式の解法を分類し述べる。

3.4.1 自己無撞着計算

コーンシャム方程式の平均場(3.32)は電子密度 $\rho(\mathbf{r})$ に依存している。しかし電子密度 $\rho(\mathbf{r})$ はコーンシャム方程式の解である一電子軌道 $\phi_i(\mathbf{r})$ のノルムの和によって表される。したがって一電子軌道を得るには平均場が必要であり、平均場を得るには一電子軌道が必要になるという循環ループに陥ることになり、コーンシャム方程式を解くことができなくなる。このような状況において方程式の解を求めるためには以下の自己無撞着計算を行うことになる。

1. まず初めに、何らかの形で電子密度を仮定する。通常は孤立原子の電子密度を初期電子配置とすることが多い。
2. 次に仮定された電子密度から平均場を計算する。
3. 計算された平均場からコーンシャム方程式を解き、一電子軌道を得る。
4. 得られた一電子軌道から電子密度を計算する。
5. 計算された電子密度と仮定した電子密度が一致すれば自己無撞着計算を終了し、得られた解からエネルギーなどの各種物理量を計算する。一致しなければ何らかの方法で電子密度を更新し、二番目のステップに戻る。

このようにコーンシャム方程式の真の解を得るためには何回も方程式を解く必要がある。このような自己無撞着計算はコーンシャム方程式に限らずハートリー法やハートリーフォック法を解く際にも使われる。電子密度の更新方法としては得られた電子密度と仮定された電子密度を混合することで新たな電子密度を生成することが多い。よく使われる手法としてブロイデン法がある。ブロイデン法はヤコビ行列の近似により準ニュートン法を解く手法であり、単純に混合するより計算の収束が速くなる。計算の収束の速さは密度汎関数計算の計算速度を左右する重要なファクターであり、他にも様々な手法が存在する。

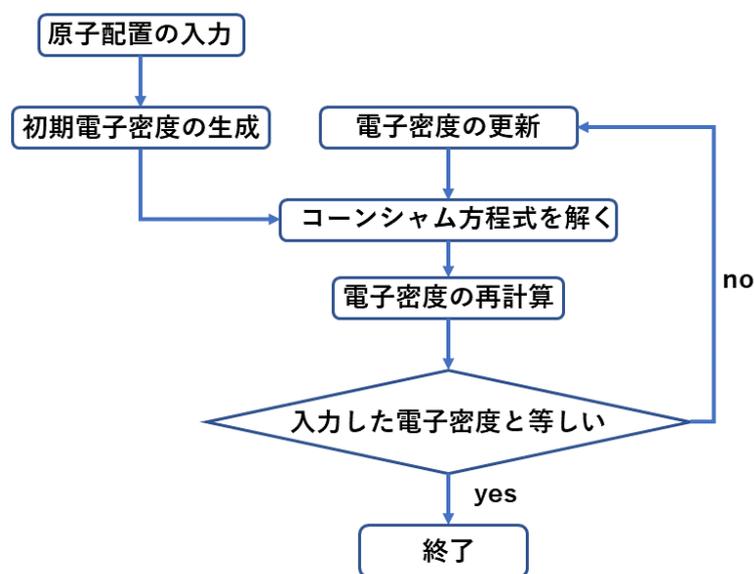


図3.1 自己無撞着計算

3.5 密度汎関数法の拡張

元来の密度汎関数法はすべての物理量を電子密度 ρ によって記述するものであったが、その後電子密度以外の物理量を基本変数に加えた理論が提案された。このような基本変数を電子密度に限らない密度汎関数法は一般化された密度汎関数法と呼ばれる。¹⁵ 基本変数に電子密度以外の変数を加えても密度汎関数法を裏付けるホーエンベルグコーンの定理は証明できる。すなわち、基本変数と外部ポテンシャルの一対一対応が保証される。そのためホーエンベルグコーンの定理からコーンシャム方程式を導いたように、新たな基本変数に対するコーンシャム方程式を導出することができる。

一般化された密度汎関数法の例としては物質の磁化 $\mathbf{m}(\mathbf{r})$ を基本変数に加えるスピン密度汎関数理論がある。¹⁶ 磁化 $\mathbf{m}(\mathbf{r})$ は密度汎関数理論の原理上、電子密度 $\rho(\mathbf{r})$ の汎関数として $\mathbf{m}(\mathbf{r})[\rho]$ 表せるはずであるがその具体的な表式は知られていない。そのため系の磁化に興味がある場合、電子密度のみを基本変数とした密度汎関数計算からは磁化の情報を知ることができない。そこで磁化 $\mathbf{m}(\mathbf{r})$ を電子密度 $\rho(\mathbf{r})$ とあわせて密度汎関数計算の基本変数とすれば、結果として基底状態の磁化 $\mathbf{m}(\mathbf{r})$ を得ることができる。実際の計算では基本変数として上向きスピンの電子密度と下向きスピンの電子密度を使うことが多い。この電子スピン密度を基本変数に加えたスピン密度汎関数法は磁気的性質の第一原理計算手法として磁性が発現している物質の計算に盛んに適用されている。

基本変数に他の物理量を加えることのもう一つの利点は、外部ポテンシャルとの作用を簡潔に書くことができることである。再びスピン密度汎関数について考えると、外部磁場 $\mathbf{B}(\mathbf{r})$ が加えられた系について密度汎関数計算を行う場合、外部磁場と磁化との相互作用は $\mathbf{m} \cdot \mathbf{B}$ と簡潔に表すことができる。基本変数に電子密度しか使わない場合は電子密度と磁場の相互作用を考えなければならないため、定式化が困難になる。このように解析の対象となる系によって適当な基本変数を選ぶことにより、通常の密度汎関数法では実行できない物理量の計算を行うことができるようになる。

基本変数に電子密度以外の物理量を加える際に起こる大きな問題は、その基本変数に対する交換相関汎関数を求めなければならないことである。新たな物理量を基本変数とする理論を構築する際は、その基本変数に対するコーンシャム方程式を導くことに加え、その基本変数に対する交換相関汎関数を計算できる形で求められなければ、実際に計算を行うことができない。交換相関汎関数の開発は密度汎関数法の高精度化に加え、新たな理論への発展においても重要なものである。今日までに開発されたスピン密度汎関数法以外の密度汎関数法としては、基本変数に常磁性電流密度 $\mathbf{j}_p(\mathbf{r})$ が加えられた電流密度汎関数理論¹⁷や、電子対密度 $\chi(\mathbf{r}, \mathbf{r}') = \langle \psi_{\uparrow}(\mathbf{r})\psi_{\downarrow}(\mathbf{r}') \rangle$ を基本変数に加えた超伝導密度汎関数理論¹⁸などが存在する。これらの理論を完成させるため基本変数に対する交換相関汎関数の開発が盛んに行われている。

第4章

交換相関汎関数

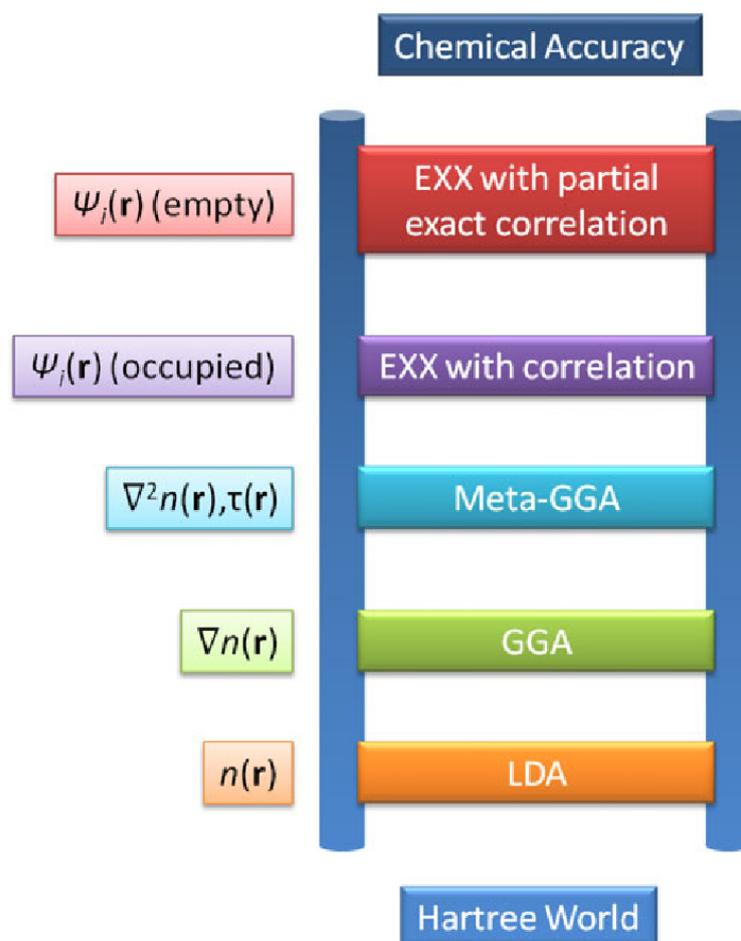
先の章ではホーエンベルグコーンの定理から密度汎関数法の基礎方程式であるコーンシャム方程式を導出した。式の導出の過程において、古典的なクーロン相互作用では捉えることのできない電子の多体効果のすべてを交換相関汎関数として表すことでコーンシャム方程式は多体シュレーディンガー方程式の厳密な一電子問題としての書き換えとなった。しかし電子の多体効果のすべてを厳密に記述することができる交換相関汎関数の表式は現在でも知られていないため、密度汎関数計算を行う際にはこの汎関数の近似形を使うことになる。密度汎関数計算における計算誤差の多くはこの交換相関汎関数の近似の不完全さに起因するものである。したがって厳密な交換相関汎関数の形が分かれば密度汎関数法は近似を含まない完全な理論となり物性物理の問題のほとんどすべてが解決できたことになる。その意味で厳密な交換相関汎関数の開発は物性物理の究極の目標であるといえる。今日まで正しい電子の多体相互作用を記述できるよう様々な交換相関汎関数が開発されてきた。この章ではそれらの汎関数を分類し、主要なものについてその特徴と問題点について述べる。

4.1 交換相関汎関数の分類

交換相関汎関数の理論のレベルは階層構造になっており、その階層構造をパーデューらは聖書に登場するヤコブの梯子にたとえた。¹⁹ ヤコブの梯子は神の使いが天界と地上界の間を行き来できるように掛けられた梯子である。交換相関汎関数における天界とは多体シュレーディンガー方程式に対する厳密な解を与える交換相関汎関数であり、地上は交換相関汎関数を無視した理論、すなわち古典的なクーロン相互作用しか含まないハートリー法である。梯子の上の方ほど電子の相互作用に関する情報をより多く含む汎関数であるが、梯子の上方の理論がいかなる系に対しても下方の理論より計算精度が良いわけではない。次節からはヤコブの梯子のそれぞれの階層の理論について紹介する。

4.1.1 局所密度近似(LDA)汎関数

交換相関汎関数におけるヤコブの梯子の一段目は局所密度近似(LDA)汎関数である。局所密度近似とはある点の交換相関エネルギーを、その点の電子密度を持つ一様電子ガスの交換相関エネルギーで近似する手法である。すな

図4.1 交換相関汎関数におけるヤコブの梯子¹⁹

わち交換相関エネルギー $E_{xc}[\rho(\mathbf{r})]$ を

$$E_{xc}^{\text{LDA}}[\rho(\mathbf{r})] = \int \varepsilon_{xc}(\rho(\mathbf{r}))\rho(\mathbf{r})d\mathbf{r} \quad (4.1)$$

と近似する。ここで $\varepsilon_{xc}(\rho(\mathbf{r}))$ は交換相関エネルギー密度である。局所密度近似から導かれる交換相関ポテンシャルは

$$V_{xc}(\mathbf{r}) = \varepsilon_{xc}(\rho(\mathbf{r})) + \rho(\mathbf{r})\frac{d\varepsilon_{xc}(\rho)}{d\rho} \quad (4.2)$$

となり、電子密度 $\rho(\mathbf{r})$ の関数となる。(4.2)から分かるように、系の中のある点の交換相関ポテンシャルを求める際は、その点の電子密度の値しか考慮されず完全に局所的である。局所密度近似は交換相関汎関数の理論では最も単純なものであるにもかかわらず、多くの系で実験結果と良い一致を示す。結晶の構造解析では格子定数を誤差1%程度で再現し、弾性定数も誤差数%程度で予測可能である^{2,12}

最も古い部類のLDA汎関数は密度汎関数法が誕生する以前にトーマスフェルミ理論を改良するためにディラックが提案した交換相互作用に関するLDA汎関数である。

$$E_x = -\frac{3}{4}\left(\frac{3}{\pi}\right)^{\frac{1}{3}}\int d^3\mathbf{r}\rho^{\frac{4}{3}}(\mathbf{r}) \quad (4.3)$$

このディラックのLDA交換汎関数は交換エネルギーの局所密度近似としての厳密解であるため,新たに開発された交換汎関数の試金石として重要な意味を持つ.

密度汎関数法以前に存在していたLDA汎関数は他にも存在する. スレーターはハートリーフォック法における交換相互作用項を局所密度近似で表した. 交換相互作用は

$$V_x(\mathbf{r}) = e^2 \sum_{j=1(N//)}^N \int \frac{\phi_j^*(\mathbf{r}')\phi_i(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|} d\mathbf{r}' \quad (4.4)$$

と表されるが,この項を計算するには各軌道ごとにその軌道と平行なスピン座標を持つ軌道との積分(交換積分)を実行せねばならず,非常に計算コストが大きくなる. そこでスレーターはこの交換相互作用項(4.4)を平均化し以下の交換汎関数を提案した.

$$V_x(\mathbf{r}) = -3e^2 \left[\frac{3}{8\pi} \rho(\mathbf{r}) \right]^{\frac{1}{3}} \quad (4.5)$$

その後平均化に伴う誤差をパラメータ α に含めた $X\alpha$ 法が提案された. $X\alpha$ 法は電子密度さえ分かれば交換エネルギーを計算できるため,それまで,固体結晶系には適用できなかったハートリーフォック法を初めて計算可能にした.

交換汎関数は局所密度極限においてその厳密系が分かっているが,相関汎関数の厳密系は未だ分かっていない.⁵ これまでの研究により電子密度の高密度極限,低密度極限に関する汎関数の表現が提案されてきたが,その中間領域,すなわち現実の物質がとり得る電子密度領域に関しては何の制限もない. そのため相関汎関数のLDA汎関数はそれらの極限における値を内挿し,その値に合うよう汎関数に含まれるパラメータをフィッティングすることによって作られる. 固体物性計算でよく用いられる局所密度近似相関汎関数としてはVWN相関汎関数²⁰やPW相関汎関数²¹などがあるが,これらの汎関数もそのようにして作られている.

4.1.2 一般化勾配近似(GGA)汎関数

ヤコブの梯子の局所密度近似よりも一段上の理論は一般化勾配近似(GGA)汎関数である. GGA汎関数は交換相関エネルギーの計算に電子密度だけでなく,電子密度の勾配の情報も含める. すなわち,交換相関汎関数を

$$E_{xc}^{GGA}[\rho(\mathbf{r})] = \int \varepsilon_{xc}(\rho(\mathbf{r}), \nabla\rho(\mathbf{r}))\rho(\mathbf{r})d\mathbf{r} \quad (4.6)$$

と電子密度とその勾配の汎関数として表す. 局所密度近似では交換相関エネルギーをその点の電子密度を持った一様電子ガスのエネルギーで近似したが,当然のことながら現実の物質では電子密度は一様でない. 汎関数に密度勾配を取り入れることは,その電子密度の局所的な変化をある程度考慮することができるようになるため,局所密度近似を改善することが期待される. GGA交換汎関数の一般的な形は

$$E_x = -\frac{1}{2} \sum_{\sigma} \int d\mathbf{r} \rho_{\sigma}^{\frac{4}{3}}(\mathbf{r}) K_{\sigma} \quad (4.7)$$

と表される. ここで σ は電子のスピン座標を表し, K_σ は無次元のパラメータ

$$x_\sigma = \frac{|\nabla\rho_\sigma|}{\rho_\sigma^{\frac{4}{3}}} \quad (4.8)$$

によって表される. これまでに数多くのGGA交換汎関数が開発されてきたが,GGA交換汎関数の違いは K_σ の x_σ に対する依存性の違いのみである. 以下では主要なGGA交換汎関数を紹介する.

B88交換汎関数²²は電子密度勾配の大きい領域の汎関数形を長距離漸近相互作用条件を満たすようにして作られた.

$$K_\sigma^{B88} = K_\sigma^{LDA} + \frac{2\zeta x_\sigma^2}{1 + 6\zeta x_\sigma \sinh^{-1} x_\sigma} \quad (4.9)$$

$$K_\sigma^{LDA} = 3 \left(\frac{3}{4\pi} \right)^{\frac{1}{3}}$$

ζ は半経験的パラメータであり, $\zeta = 0.0042$ という値が使われる. B88交換汎関数は交換エネルギーをそれまで開発されてきた交換汎関数と比べて格段に精度良く再現することに成功し,今もなお密度汎関数計算で最もよく使われる交換汎関数の一つになっている. B88交換汎関数は長距離漸近相互作用条件を満たすように作られた汎関数であるため,それ以外の汎関数に対する基礎物理条件をそれほど満たさない.

PW91交換汎関数²³はB88交換汎関数をより多くの基礎物理条件を満たすように改良して作られた交換汎関数であり,多くのパラメータを含む複雑な表式になっている. PW91汎関数は多くの基礎物理条件を満たすが, x_σ が大きい場合に K_σ が急速に0に近づき計算に悪影響を与えることが問題となっている.

PBE交換汎関数²⁴はPW91交換汎関数を満たすべき基礎物理条件を重要なものに限定し,パラメータを二つまでに減らして簡略化することで作られた.

$$K_\sigma^{PBE} = K_\sigma^{LDA} \left[1 + \kappa - \frac{\kappa}{1 + \frac{\mu x_\sigma^2}{(48\pi^2)^{\frac{2}{3}} \kappa}} \right] \quad (4.10)$$

PBE交換汎関数の K_σ は x_σ が大きい領域では一定値に近づくためPW91交換汎関数における計算の不安定性の問題を解消する. PBE交換汎関数は多くの系で高い計算精度を誇り,固体物理の分野でよく使用される汎関数の一つである.

GGA相関汎関数は局所密度近似の場合と同様に,相関エネルギーに対する基礎物理条件を満たすように定式化を行い,それらの条件に合わせてパラメータをフィッティングすることで開発される.

PW91相関汎関数は基礎物理条件をなるべく多く満たすようにして作られた相関汎関数であり11個ものパラメータを持っている. PW91相関汎関数は多くの基礎物理条件を満たすが,そのパラメータの多さに起因する問題を抱

えている。

PBE相関汎関数はPW91相関汎関数の問題を解決するために作られた相関汎関数であり,パラメータ数を二個まで減らし,重要な基礎物理条件のみを満たすように作られた。

$$\begin{aligned}
 E_c^{PBE} &= E_c^{PW-LDA}[\rho] + \int d\mathbf{r} \rho H[\rho, \zeta, t] \\
 H[\rho, \zeta, t] &= \gamma \phi^3 \ln \left[1 + \frac{\beta}{\gamma} t'^2 \left(\frac{1 + At'^2}{1 + At'^2 + A^2 t'^4} \right) \right] \\
 A &= \frac{\beta}{\gamma} \left[\exp \left(-\frac{E_c^{PW-LDA}[\rho]}{\gamma \phi^2 \rho} \right) - 1 \right]^{-1} \\
 \phi &= \frac{1}{2} \left[(1 + \zeta)^{\frac{2}{3}} + (1 - \zeta)^{\frac{2}{3}} \right]
 \end{aligned} \tag{4.11}$$

PBE相関汎関数はパラメータを γ, β の二個しか含まずPW91相関汎関数より優れた汎関数である。本研究では,このPBE交換相関汎関数を雛型とし,新たな交換相関汎関数の開発を行った。

4.1.3 meta GGA汎関数

ヤコブの梯子の三段目の理論としてmetaGGA汎関数がある。metaGGA汎関数は交換相関汎関数に電子密度 $\rho(\mathbf{r})$ とその一次勾配 $\nabla\rho(\mathbf{r})$ だけでなく,電子の運動エネルギー密度 τ を含む。電子密度の二次勾配 $\nabla^2\rho(\mathbf{r})$ でなく,運動エネルギー密度 τ を加える理由は,電子密度の二次勾配 $\nabla^2\rho(\mathbf{r})$ は全空間で積分すると0になるほど効果が小さく,数値的に不安定だからである。電子の運動エネルギー密度 τ は

$$\begin{aligned}
 \tau &= \sum_{\sigma} \tau_{\sigma} = \frac{1}{2} \sum_{\sigma} \sum_i |\nabla \phi_{i\sigma}|^2 \\
 &= \sum_{\sigma} \left[C_F \rho_{\sigma}^{\frac{5}{3}} + \frac{1}{72} \frac{|\nabla \rho_{\sigma}|^2}{\rho_{\sigma}} + \frac{1}{6} \nabla^2 \rho_{\sigma} + \mathcal{O}(\nabla^4) \right]
 \end{aligned} \tag{4.12}$$

と展開することができ,電子密度の二次勾配 $\nabla^2\rho(\mathbf{r})$ の情報が含まれていることが分かる。代表的なmetaGGA交換相関汎関数としてはPKZBmeta交換相関汎関数²⁵がある。PKZBmeta交換相関汎関数はPBE交換相関汎関数を物理的背景に持ち,運動エネルギー密度を用いて拡張したものである。それによりPBE交換相関汎関数より高精度な計算を行うことができる。他にもTPSSmeta交換相関汎関数やFT98meta交換相関汎関数,KCISmeta交換相関汎関数などがある。

4.1.4 混成汎関数

混成汎関数はハートリーフォック法で登場する交換相互作用項を用いる汎関数であり,ヤコブの梯子の四段目にあたる。一般的な混成汎関数は既存の交換汎関数とハートリーフォック法の交換相互作用項の線形和をとるが,これは既存の交換汎関数をハートリーフォック法の交換相互作用項で補正するというよりは,ハートリーフォック法の交換相互作用とGGA交換汎関数の間に交換汎関数の厳密解が存在しているというansatzが設計思想の基になっている。混成汎関数の中で最も有名で利用されてきたものはB3LYP混成汎関数²⁶である。この汎関数はハートリー

フォック交換相互作用項に加え,B88交換汎関数,LYP相関汎関数,LDA交換相関汎関数の線形結合により構成されており,3個のパラメータを含む.

$$E_{xc}^{B3LYP} = E_{xc}^{LDA} + a_1(E_x^{HF} - E_x^{LDA}) + a_2\Delta E_x^{B88} + a_3(E_c^{LYP} - E_c^{VWN-LDA}) \quad (4.13)$$

この汎関数のパラメータ a_1, a_2, a_3 は数十種類の原子や分子からなるG2ベンチマークセットの物性値を再現するように決められた. B3LYP汎関数は小分子の物性値を高い精度で再現することができるが,化学反応計算や大規模分子系において様々な問題が指摘されている. B3LYP以外にも様々な混成汎関数が存在するが,それらについては次章の自己相互作用誤差の節で紹介する.

4.1.5 半経験的汎関数

半経験的汎関数はヤコブの梯子には含まれていないが,交換相関汎関数の設計指針の一つとして本研究との関連があるため,ここで紹介する. 半経験的汎関数とは汎関数に含まれるパラメータを無制限に使うことで,物性値を再現できるようにした汎関数である. 最初に登場した半経験的汎関数はB97汎関数²⁷であり,現在の半経験的汎関数の基礎となっている. この汎関数はLDA交換相関汎関数とハートリーフォック交換相互作用項からなっており,汎関数に含まれるパラメータは13個である. Mx系汎関数^{28,29}はハートリーフォック交換相互作用項とmetaGGA汎関数を含む混成metaGGA汎関数であり,38個ものパラメータを含む. これら以外にも様々な半経験的汎関数が存在しており,物性を高い精度で再現することができるが,全く再現できない化学反応が存在も報告されている. また,汎関数を構成する交換相関相互作用項の依存性が不明確であることなどから計算結果の妥当性を保証することが難しいという問題がある.

4.2 交換相関汎関数に対する要請

交換相関汎関数の良さは様々な系の物性値をどれだけ精度良く再現できるかということに加え,多体電子系が満たすべき基礎的な物理条件をどれほど満足するかということによって評価される. 先の節で見てきた汎関数もその開発の起点になるのは,この基礎物理条件を満たす表式的设计であることが多い. この節では交換相関汎関数が満たすべき基礎物理条件について主要なものをまとめる.

負値条件

電子の運動,交換,相関エネルギーの符号は,0でない電荷密度 ρ に対して,

$$\begin{aligned} T[\rho] &> 0 \\ E_x[\rho] &< 0 \\ E_c[\rho] &\leq 0 \end{aligned} \quad (4.14)$$

が常に成立する. 運動,交換エネルギーは電子が存在する限り0にはならないが,相関エネルギーは1電子系では0となる.

局所密度近似極限条件,一般化勾配近似極限条件

電子密度勾配 $\nabla\rho$ が電子密度 ρ に比べて十分小さい極限と、その逆の $\nabla\rho$ が ρ に比べ十分大きい極限における運動,交換,相関エネルギーに関する条件である。前者の条件は運動,交換,相関エネルギーに関するものであるが,後者は運動,相関エネルギーについてのものであり,交換エネルギーに関しては極限条件が見つからない。

座標スケール条件

座標の単位を極限まで大きくあるいは小さくした場合に対応する,運動,交換,相関エネルギーの振る舞いを規定する条件である。スケール条件の種類は x, y, z 座標すべてを拡大,縮小する均一座標スケール条件, x, y 座標のみを拡大,縮小する二次不均一座標スケール条件, x 座標のみを拡大,縮小する一次不均一座標スケール条件がある。

これらの基礎物理条件以外にもリーブ・オクスフォード束縛条件,一電子自己相互作用無誤差条件,長距離漸近相互作用条件などがある。これまでに紹介してきた交換相関汎関数がこれらの基礎物理条件をどの程度満たすかを示したのが図4.2,4.3である。⁵

条件	LDA	PW91	PBE
負値条件	○	○	○
局所密度近似極限条件	○	○	○
一般化勾配近似極限条件	-	○	○
三次均一座標スケール条件	○	○	○
二次不均一座標スケール条件	×	×	×
一次不均一座標スケール条件	×	×	×
リーブ・オクスフォード束縛条件	○	○	○
一電子自己相互作用無誤差条件	×	×	×
長距離(遠隔)漸近相互作用条件	×	×	×

図4.2 交換汎関数に対する基礎物理条件と各交換汎関数の可否⁵

条件	LDA	PW91	PBE
負値条件	○	×	○
局所密度近似極限条件	×	×	×
一般化勾配近似極限条件	-	○	○
高密度勾配極限	×	○	○
三次均一座標スケール条件	×	×	○
二次不均一座標スケール条件	×	○	×
一次不均一座標スケール条件	×	○	×
一電子自己相互作用無誤差条件	×	×	×

図4.3 相関汎関数に対する基礎物理条件と各相関汎関数の可否⁵

すべての基礎物理条件を満たす交換相関汎関数は存在していない。厳密な交換相関汎関数を得るにはこれらの条件をすべて満たし、かつあらゆる物性値を満足するような汎関数を設計しなければならない。基礎物理条件をすべて満たすような汎関数の表式を設計するだけでも至難の業であるといえる。

第5章

バンドギャップ

バンドギャップは物性値の中でも特に重要な量である。物質の電気伝導はバンドギャップによって支配されており、その値が0であれば金属、室温程度のエネルギースケールであれば半導体、それより大きければ絶縁体となる。また光学素子においてもバンドギャップはその発光、吸光特性を決めており、望ましい光学素子を設計するにはバンドギャップを制御することが必須である。現在、超格子や量子ドットなどのナノテクノロジーによってバンドギャップを人工的に制御するバンドエンジニアリング注目を集め、盛んに研究されている。これらの物性研究の場では、バンドギャップの値を第一原理的に計算し予測することが求められるが、このバンドギャップこそ密度汎関数法が抱える最大の問題の一つであり、交換相関汎関数の不完全さが顕著に表れる量である。例えばSiのバンドギャップをLDA汎関数によって計算すると、実験値の1/2~1/3と計算される。このバンドギャップの過小評価はGGA汎関数にも共通している。図5.1は様々な物質に対しGGA汎関数であるPBE交換相関汎関数を用いバンドギャップの計算を行った結果である。³⁰多くのバンドギャップの値が過小評価されている。

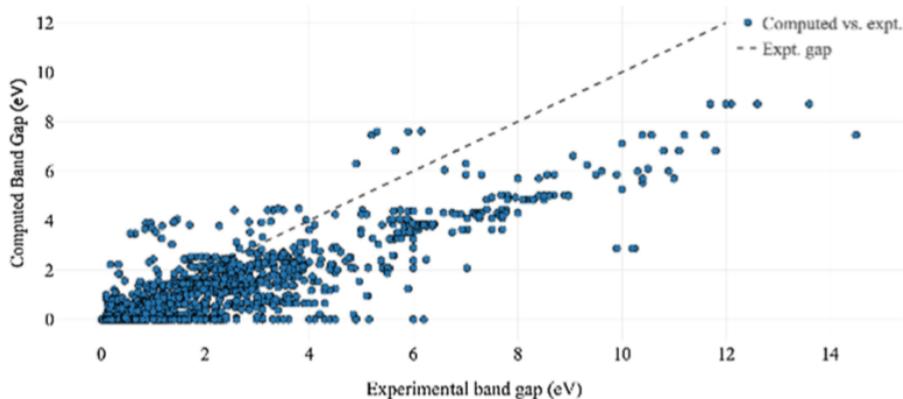


図5.1 PBE交換相関汎関数によりバンドギャップを計算した結果³⁰

このバンドギャップの問題を解決するため現在に至るまで様々な研究がなされてきた。この章では密度汎関数法が抱えるバンドギャップの問題の原因と、その改善法について述べる。

5.1 ヤナックの定理と $N \pm 1$ 電子系

密度汎関数法で計算されるバンドギャップはコーンシャム方程式を解いて得られる一電子軌道（コーンシャム軌道）の最低非占有軌道のエネルギー準位 ε_{N+1} と最高占有軌道のエネルギー準位 ε_N の差

$$E_g^{\text{KS}} = \varepsilon_{N+1} - \varepsilon_N \quad (5.1)$$

として計算される。このことはバンド理論から考えれば妥当に思われるが、コーンシャム方程式から導かれるエネルギー固有値（コーンシャム準位）が本当に物質中の電子のエネルギー準位と対応しているかは明らかではない。このコーンシャム準位の意味はヤナックの定理によって与えられる。ヤナックの定理は系の全体のエネルギー E のコーンシャム軌道 ϕ_i における電子の占有数 n_i に関する微分値がコーンシャム準位 ε_i に対応するという定理である。通常電子のそれぞれの軌道に対する占有数は0か1かであるが、ヤナックの定理では占有数を非整数に拡張する。このような自然数でない占有数は分数占有系と呼ばれる。分数占有系の電子密度 $\rho(\mathbf{r})$ と運動エネルギー $T_s[\rho]$ は

$$\begin{aligned} \rho(\mathbf{r}) &= \sum_i n_i |\phi_i(\mathbf{r})|^2 \\ T_s &= \sum_{i=1}^N n_i \int \phi_i^*(\mathbf{r}) \left(-\frac{\hbar^2}{2m} \nabla^2 \right) \phi_i(\mathbf{r}) d\mathbf{r} \\ &= \sum_{i=1}^N n_i \varepsilon_i - \int V_{\text{eff}}(\mathbf{r}) \rho(\mathbf{r}) d\mathbf{r} \end{aligned} \quad (5.2)$$

と定義される。電子密度 $\rho(\mathbf{r})$ を占有数 n_i で微分すると、

$$\frac{\partial \rho(\mathbf{r})}{\partial n_i} = |\phi_i(\mathbf{r})|^2 \quad (5.3)$$

となる。また運動エネルギー T_s の微分は

$$\frac{\partial T_s[\rho]}{\partial \rho(\mathbf{r})} = \varepsilon_i - V_{\text{eff}}(\mathbf{r}) |\phi_i(\mathbf{r})|^2 \quad (5.4)$$

となる。全エネルギー E は

$$\begin{aligned} E &= F[\rho(\mathbf{r})] + \int V(\mathbf{r}) \rho(\mathbf{r}) d\mathbf{r} \\ &= T_s[\rho(\mathbf{r})] + \frac{e^2}{2} \int \frac{\rho(\mathbf{r}) \rho(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r} d\mathbf{r}' + E_{\text{xc}}[\rho(\mathbf{r})] \end{aligned} \quad (5.5)$$

となるので、これを占有数 n_i で微分すると、

$$\frac{\partial E}{\partial n_i} = \varepsilon_i \quad (5.6)$$

というヤナックの定理が導かれる。このヤナックの定理を使って密度汎関数法におけるバンドギャップについて考察していく。

バンドギャップは基底状態のイオン化エネルギー I と電子親和力 A を用いて、

$$E_g = I - A = \{E(N-1) - E(N)\} - \{E(N) - E(N+1)\} \quad (5.7)$$

と定義される. ここで N 電子系の基底状態のエネルギーを $E(N)$ とした. ヤナックの定理(5.6)を使うと, 電子のイオン化エネルギー I , 電子親和力 A はそれぞれ,

$$\begin{aligned} I &= E(N-1) - E(N) = - \int_0^1 dn_N \varepsilon_N(N-1+n_N) \\ A &= E(N) - E(N+1) = - \int_0^1 dn_{N+1} \varepsilon_{N+1}(N+n_{N+1}) \end{aligned} \quad (5.8)$$

となる. 分数占有状態 $N+n_{N+1}$ 電子系とは N 電子系と $N+1$ 電子系が n_{N+1} の割合で混合したものであると考えることができるべきであるので,¹² $N+n_{N+1}$ 電子系のエネルギー $E(N+n_{N+1})$ は

$$E(N+n_{N+1}) = (1-n_{N+1})E(N) + n_{N+1}E(N+1) \quad (5.9)$$

と書かれるべきである. $E(N-1+n_N)$ についても同様である. これにさらにヤナックの定理を使うと,

$$\begin{aligned} \varepsilon_{N+1}(N+n_{N+1}) &= -E(N) + E(N+1) \\ \varepsilon_N(N-1+n_N) &= -E(N-1) + E(N) \end{aligned} \quad (5.10)$$

となる. 右辺は n_i に依存していないため, 左辺はそれぞれ $\varepsilon_{N+1}(N+1)$, $\varepsilon_N(N)$ に等しい. これより, バンドギャップは

$$E_g = \varepsilon_{N+1}(N+1) - \varepsilon_N(N) \quad (5.11)$$

と表すことができる. すなわちバンドギャップは $N+1$ 電子系の $N+1$ 番目のエネルギー準位と N 電子系の N 番目のエネルギー準位の差として定義される. 一方コーンシャム方程式を解いて計算されるバンドギャップ E_g^{KS} は

$$E_g^{KS} = \varepsilon_{N+1}(N) - \varepsilon_N(N) \quad (5.12)$$

であり, 厳密に定義されたバンドギャップのものとは異なる. 厳密なバンドギャップの表式はエネルギーの線形依存性(5.9)とヤナックの定理を使うことで E_g^{KS} と

$$E_g = E_g^{KS} + \left. \frac{\delta E_{xc}}{\delta \rho} \right|_{M=N-0}^{M=N+0} \quad (5.13)$$

という関係にあることを導くことができる. M は非整数の電子数である. この式から, 厳密なバンドギャップとコーンシャムギャップの間のずれは交換相関ポテンシャルが電子数 M が整数の時に不連続に変化することに起因する跳びによるものであることが分かる.³¹ 一次元系の数値厳密解を調べた結果, この跳びはモット絶縁体や電荷密度波状態が基底状態になる場合は無視できないが, 半導体の場合その値は小さいことが分かっている. したがって, 密度汎関数法におけるバンドギャップの計算誤差の原因は交換相関ポテンシャルの不連続性によるものと, 交換相関ポテンシャルの近似誤差によるものが複合的に寄与している.²

5.2 自己相互作用

密度汎関数法におけるバンドギャップの問題の原因は先の節で論じた交換相関汎関数の不連続性に起因するものもあるが, 最大の原因は自己相互作用である. 本論文では節を分けたが, 自己相互作用は交換相関汎関数の不連続性にもかかわる問題であるので, 両者を独立した問題として考えることは不適當である. 自己相互作用を考えるの

に最適なのは一電子系である.ここでは水素原子を例に自己相互作用について考える.水素原子には一つの電子しか含まれておらず,他の電子とのクーロン相互作用は生じない.ここでハートリー法の平均場

$$\left\{ -\sum_i \frac{\hbar^2}{2m} \nabla_i^2 + \sum_i v_{\text{ext}}(\mathbf{r}_i) + e^2 \sum_{j \neq i} \int \frac{|\phi_j(\mathbf{r}')|^2}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r}' \right\} \phi_i(\mathbf{r}) = \epsilon_i \phi_i(\mathbf{r}) \quad (5.14)$$

について見てみると,ハミルトニアン第三項は平均的な電子密度とのクーロン相互作用を表している.水素原子に含まれる一つの電子もそれ自身で電子密度分布を形成しているため,水素原子にハートリー法を適用すると,電子は自分自身が作り出した電子密度と現実には存在しないクーロン相互作用をすることになる.この電子自身が作り出した電子密度との非物理的な相互作用を自己相互作用という.

ここでハートリーフォック法の平均場を見てみると,

$$\left\{ -\sum_i \frac{\hbar^2}{2m} \nabla_i^2 + \sum_i v_{\text{ext}}(\mathbf{r}_i) + e^2 \sum_{j=1}^N \int \frac{|\phi_j(\mathbf{r}')|^2}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r}' - e^2 \sum_{j=1(j \neq i)}^N \int \frac{\phi_j^*(\mathbf{r}') \phi_i(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r}' \right\} \phi_i(\mathbf{r}) = \epsilon_i \phi_i(\mathbf{r}) \quad (5.15)$$

交換相互作用項の $i = j$ 成分はハートリー項と符号が反対になっている.そのためハートリー項と交換相互作用項の和をとれば自己相互作用を打ち消すことができる.このことは密度汎関数法におけるバンドギャップの改善方法の指針として重要な事実である.次の節ではこのことを踏まえバンドギャップをどのようにして改善するかについて述べる.

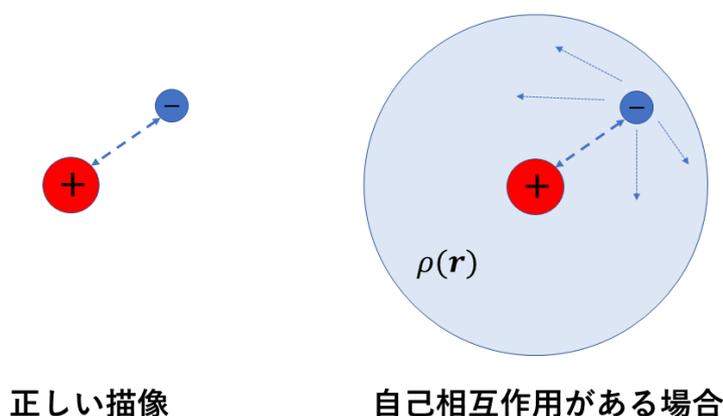


図5.2 自己相互作用

5.3 ギャップ誤差の改善法

先の節までの議論から密度汎関数法におけるギャップ誤差の原因は交換相関汎関数の不連続性と自己相互作用によるものがあることが分かった.この節ではこのことを踏まえてギャップ誤差を改善する手法としてよく使われるものを紹介する.

5.3.1 混成汎関数

自己相互作用の節でも述べたが,自己相互作用はハートリー項に交換相互作用項を加えることで打ち消すことができる. このようにハートリーフォック法の交換相互作用項を用いて作られた交換相関汎関数は先の章で登場した混成汎関数である. 実際,混成汎関数はギャップ誤差を大幅に改善することが報告されている. ここではギャップ誤差の改善で成果をあげたPBE0混成汎関数³²と,HSE混成汎関数³³について紹介する. PBE0混成汎関数はPBE交換相関汎関数を拡張し,PBE汎関数の交換エネルギーをハートリーフォック交換相互作用項で置き換えたものである.

$$E_{xc}^{PBE0} = E_{xc}^{PBE} + \frac{1}{4} (E_x^{HF} - E_x^{PBE}) \quad (5.16)$$

この汎関数は経験的なパラメータを含まないが,バンドギャップをはじめ物性値を高い精度で予測することが可能である.

HSE混成汎関数はGGA交換汎関数にハートリーフォック交換相互作用項の短距離部分のみを混合した汎関数であり,バンドギャップを高精度に再現する.

これらの汎関数を用いてバンドギャップの計算を行った結果が図5.3である.³⁴

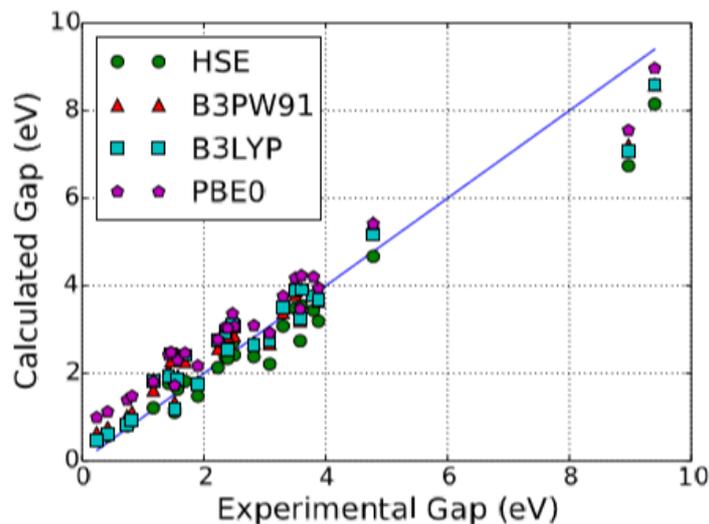


図5.3 混成汎関数によりバンドギャップの予測を行った結果³⁴

どちらの汎関数も高い精度でバンドギャップを予測することができている.

5.3.2 GW近似

GW近似は密度汎関数法の枠組みには入らないが,バンドギャップを補正する方法として大変有用であるためここで取り上げる. 固体中の電子はその周りに正の電荷を帯びた準粒子として存在しており,クーロン相互作用がある程度遮蔽されている. ハートリーフォック法を固体のバンド計算に用いるとバンドギャップを過大評価するが,それはこの遮蔽効果を無視したことが原因である. そこでこの遮蔽されたクーロン相互作用 $W = \epsilon^{-1}v$ を電子系の

計算に取り入れることができれば、より正しく多体電子系の描像を捉えることができると期待される。ここで ϵ^{-1} は誘電関数、 v は裸のクーロン相互作用である。ヘディンはグリーン関数 G 、分極率 P 、遮蔽クーロン相互作用 W 、自己エネルギー Σ の間に成り立つ自己無撞着な関係式を導出した。³⁵ これらの物理量を自己無撞着計算により解けば自己エネルギー Σ から、コーンシャム準位のエネルギーを

$$\epsilon_i = \epsilon_i^{\text{KS}} + \int \int \phi_i^*(\mathbf{r}) (\Sigma(\mathbf{r}, \mathbf{r}'; \epsilon_i) - V_{\text{xc}} \delta(\mathbf{r} - \mathbf{r}')) \phi_i(\mathbf{r}') d\mathbf{r} d\mathbf{r}' \quad (5.17)$$

と補正することで、固体中の準粒子としての電子のエネルギー準位を求めることができる。ヘディンによって求められた関係式は自己無撞着型のものであったが、これを近似して、自己エネルギーを

$$\Sigma^{\text{GW}}(\mathbf{r}, \mathbf{r}'; t) = iG(\mathbf{r}, \mathbf{r}'; t)W(\mathbf{r}, \mathbf{r}'; t) \quad (5.18)$$

としたのがGW近似である。これを固体のバンドギャップの計算に用いた例が図5.4である。³⁶

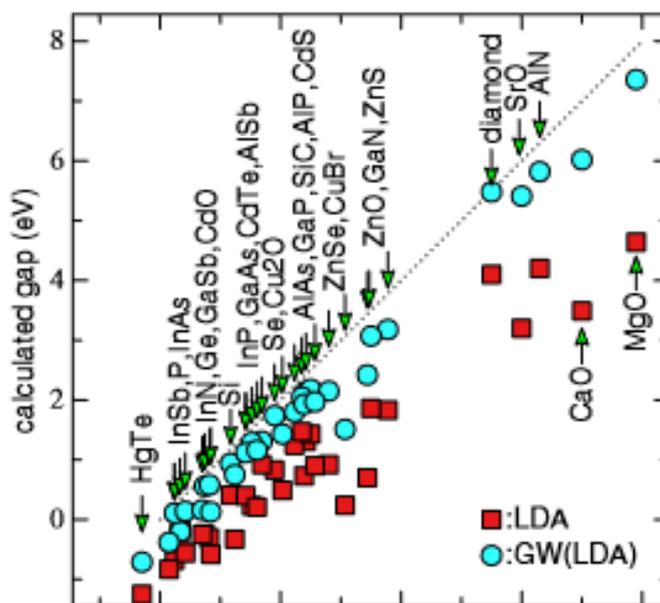


図5.4 GW近似によりバンドギャップの予測を行った結果³⁶

GW近似によりLDAで計算されたギャップ誤差を大幅に改善することに成功している。

密度汎関数法とGW近似の関りは、GW近似を行うには電子のグリーン関数を求める必要があるが、このグリーン関数を構成する際に密度汎関数計算で求められたコーンシャム軌道が使われることが多い。

5.3.3 DFT+U

DFT+Uはd軌道やf軌道が部分的に占有された遷移金属酸化物などの局在電子系におけるバンドギャップの補正に使われる。電子は通常、固体中を特定の波数を持った波として遍歴しているが、電子同士のクーロン相互作用が強まると互いに身動きが取れなくなり局在化する。すなわち絶縁体になる。このような電子相関が顕著に表れる系

を強相関電子系といい現在の物性物理の最先端の話題である。通常の交換相関汎関数で強相関電子系の計算を行うと、電子相関を十分に捉えきれないため実際は絶縁体であるにもかかわらずバンドギャップが消滅してしまうことが多い。この問題を解決するために、LDAなどの既存の交換相関汎関数にd軌道やf軌道におけるクーロン反発力を U というパラメータで露に取り入れることで、電子相関を再現しようとしたのがDFT+ U である。DFT+ U により消滅してしまったバンドギャップを開かせることができるが、このクーロン反発力のパラメータ U は実際のバンドギャップなどの物性値を再現できるように決められるため、もはや第一原理的ではなく普遍的なバンドギャップの改善法であるとは言えない。しかしながら強相関電子系の電子の描像を知る手段としては有用な方法であり、様々な系で適用されている。図5.5は Ce_2O_3 に対し、LDA交換相関汎関数にクーロン反発力 U を加えバンドギャップを計算した結果である。³⁷

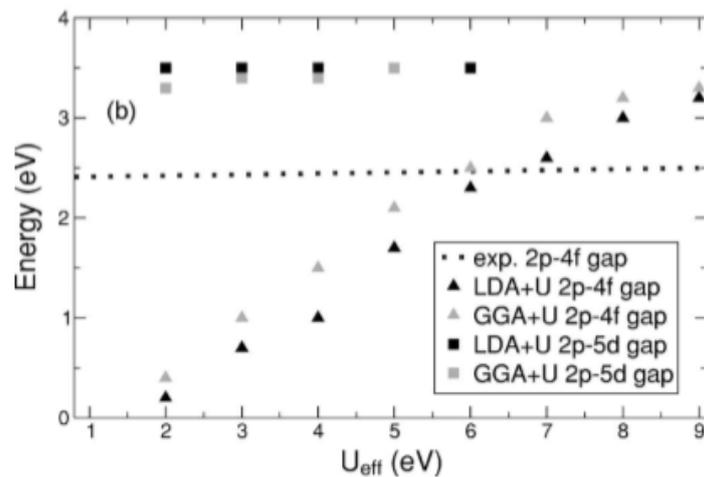


図5.5 Ce_2O_3 におけるDFT+ U の U とバンドギャップの値の依存性³⁷

U の増加に伴いバンドギャップが広がっていくことが見て取れる。

第6章

ニューラルネットワーク

本研究では三次元畳み込みニューラルネットワーク(3DCNN)を用いて密度汎関数理論の改善を試みた。この章ではその基礎となるニューラルネットワークの基本的な仕組みとについて紹介する。また、本研究の関連研究として畳み込みニューラルネットワークを物性計算に応用した手法を紹介する

6.1 ニューラルネットワークの基本構造

ニューラルネットワークは脳の構造を模倣した計算アルゴリズムである。脳は膨大なニューロンが互いに結合して構成されており、ニューロンは他のニューロンから刺激を受けると細胞体の電位が上昇していきある閾値を超えると発火しインパルス信号を発する。これを計算アルゴリズムとして初めて実装したのがローゼンブラッドのパーセプトロンである。パーセプトロンは図6.1のような単純な線形識別モデルである..

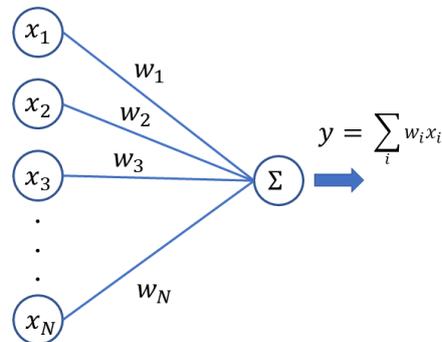


図6.1 パーセプトロンの模式図

入力信号 x_i は結合荷重 w_i で重み付けられ出力ニューロンに伝えられる。出力ニューロンは入力が0以上ならば1を、それ以下であれば-1を返す。これを数式で表すと、

$$y = f\left(\sum_i w_i x_i\right)$$

$$f(a) = \begin{cases} 1 & (a \geq 0) \\ -1 & (a < 0) \end{cases} \quad (6.1)$$

となる。結合荷重 w_i を識別誤差が減少するように更新していくことで、入力信号から望まれる出力を返すモデル

を学習することができる。パーセプトロンは第一次人工知能ブームに盛んに研究されたが、その後マービンミンスキーによってパーセプトロンが線形な分離平面を持つ識別問題でしか解を学習することができないという重大な欠陥を指摘したことで研究は下火になっていった。線形な分離平面を持たない代表的な例としてはXOR問題があるが、パーセプトロンはこの単純な問題すら解くことができない。

その後ラメルハートらはパーセプトロンの入力層と出力層の間に新たに中間層を加えた多層パーセプトロンを開発しこの問題を解決した。この多層パーセプトロンは中間層が一層であっても非常に高い汎用性を持ち、もし無限の中間ノードを持つことができれば、あらゆる連続関数を近似することができることが証明されている。このことから多層パーセプトロン、すなわち今日の人工知能研究の主役であるニューラルネットワークは万能近似器 (universal approximation) であると言われている。ここではもっとも単純な例として一層の中間層を持つニューラルネットワーク(6.2)を考える。

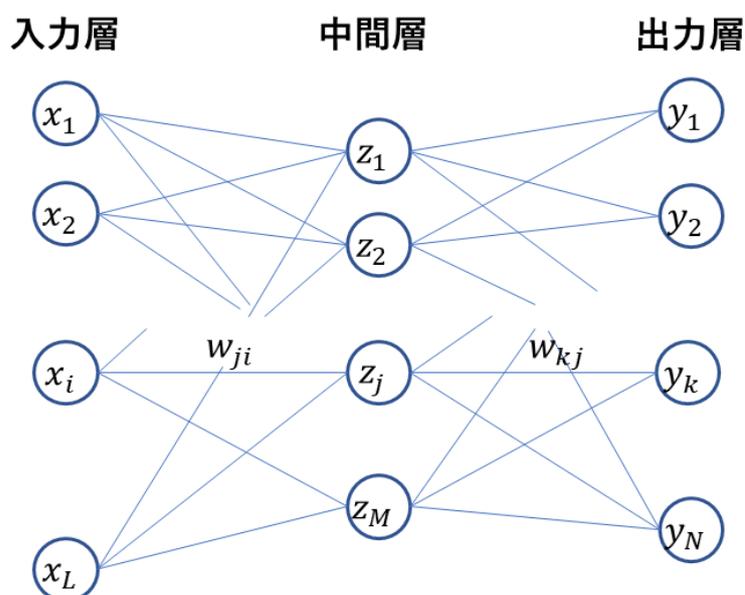


図6.2 ニューラルネットワークの模式図

ネットワークの第一層において入力変数 x_i と重みパラメータ $w_{ji}^{(1)}$ の線形和 a_j が計算される。

$$a_j = \sum_{i=1}^D w_{ji}^{(1)} x_i + w_{j0}^{(1)} \quad (6.2)$$

そして a_j の非線形関数 h による写像が中間ノード z_j の入力となる。

$$z_j = h(a_j) \quad (6.3)$$

ここでこの非線形関数 h は活性化関数と呼ばれる。この非線形な活性化関数によりニューラルネットワークは高い

表現力を獲得することになる。中間層の入力 z_j は再度重みパラメータ w_{kj} と線形和により

$$a_k = \sum_{j=1}^M w_{kj}^{(2)} z_j + w_{k0}^{(2)} \quad (6.4)$$

と変形され、出力層の活性化関数 σ を通して、出力ユニット y_k に伝えられる。

$$y_k = \sigma a_k \quad (6.5)$$

出力層の活性化関数 σ は問題によって使い分けられ、識別問題ではロジスティックシグモイド関数が、回帰問題では恒等関数などが使われる。以上をまとめるとネットワークの出力は

$$y_k(\mathbf{x}, \mathbf{w}) = \sigma \left(\sum_{j=0}^M w_{kj}^{(2)} h \left(\sum_{i=0}^D w_{ji}^{(1)} x_i \right) \right) \quad (6.6)$$

となる。ここで先ほどの単層パーセプトロンと比較すると、ニューラルネットワークは単層パーセプトロンに変数 x_i の非線形写像 z_j を入力したものであるとみなすことができる。中間層の役割はパーセプトロンでは分離することができなかつた変数の空間を、線形分離可能な別の空間へと写像することである。ネットワークの重みパラメータ \mathbf{w} を出力誤差が減少するように更新していくことで学習が行われるが、その計算は誤差逆伝播法を用いることで効率よく行うことができる。次節ではこの誤差逆伝播法について紹介し、さらにその手法を応用してネットワークの微分に関する情報を効率よく計算する方法について述べる。

6.2 誤差逆伝播法

ニューラルネットワークの出力誤差 $E(\mathbf{w})$ は一般に二乗和誤差を考えることで、

$$E(\mathbf{w}) = \frac{1}{2} \sum_{n=1}^N \{y(\mathbf{x}_n, \mathbf{w}) - t_n\}^2 \quad (6.7)$$

となる。ニューラルネットワークの学習はこの誤差関数 $E(\mathbf{w})$ を最小化する重みパラメータ \mathbf{w} を探索する最適化問題であると考えることができる。最適化問題を解くために用いられる一般的な手法としては勾配法がある。勾配法では、誤差関数 $E(\mathbf{w})$ のパラメータに関する微分 $\nabla E(\mathbf{w})$ の情報から、パラメータを

$$\mathbf{w}^{t+1} = \mathbf{w}^t - \eta \nabla E(\mathbf{w}) \quad (6.8)$$

と更新していくことで誤差関数の極小点を見つけ出す手法である。このネットワークの勾配情報を効率よく計算できる手法が誤差逆伝播法である。誤差関数 $E(\mathbf{w})$ の第一層の重み w_{ji} に関する微分は連鎖律により

$$\frac{\partial E(\mathbf{w})}{\partial w_{ji}} = \frac{\partial E}{\partial a_j} \frac{\partial a_j}{\partial w_{ji}} \quad (6.9)$$

と書くことができる。ここで誤差 δ_j を

$$\delta_j = \frac{\partial E(\mathbf{w})}{\partial a_j} \quad (6.10)$$

と定義すると

$$\frac{\partial a_j}{\partial w_{ji}} = z_i \quad (6.11)$$

より,誤差関数の微分は

$$\frac{\partial E(\mathbf{w})}{\partial w_{ji}} = \delta_j z_i \quad (6.12)$$

と表すことができる。 z_j はネットワークの出力を計算する際に手に入れているので,誤差 δ_j が分かれば,誤差関数の微分を計算することができる。誤差 δ_j は再び微分の連鎖律を用いることで,

$$\delta_j = \frac{\partial E(\mathbf{w})}{\partial a_j} = \sum_k \frac{\partial E}{\partial a_k} \frac{\partial a_k}{\partial a_j} = h'(a_j) \sum_k w_{kj} \delta_k \quad (6.13)$$

と表される。ここで誤差の表式を見てみると,一層目の誤差 δ_j は二層目の誤差 δ_k の線形和で表されていることが分かる。これは入力から出力を得るために行う演算(順伝播)のちょうど逆の操作であり,この手法が誤差逆伝播法と呼ばれる所以である。誤差逆伝播法の最大の利点はその計算効率の良さである。もし数値微分によりネットワークの微分を求めた場合,その計算量は重みパラメータの数を W とすると, $O(W^2)$ となるが,誤差逆伝播法の場合,その計算は出力を求める演算とほぼ同じものであるため $O(W)$ で済み,非常に効率的である。またニューラルネットワークのほぼすべての演算を積和演算で行うことができるため,GPUやFMAによって計算を加速することができる。

6.2.1 ヤコビ行列

誤差逆伝播法を応用すると,誤差関数の微分値以外のネットワークの微分に関する情報も効率的に計算することができる。ここでは本研究でも用いられているネットワークのヤコビ行列の誤差逆伝播による計算法について見ていく。ヤコビ行列はネットワークの入力に対する感度の指標であるため,ネットワークの出力の不変性や頑健性を向上させるために使用されることがある。ヤコビ行列 J の各要素はネットワークの出力 y_k の入力 x_i に関する微分で与えられる。

$$J_{ki} = \frac{\partial y_k}{\partial x_i} \quad (6.14)$$

ここで,微分の連鎖律により,

$$\begin{aligned} J_{ki} &= \frac{\partial y_k}{\partial x_i} = \sum_j \frac{\partial y_k}{\partial a_j} \frac{\partial a_j}{\partial x_i} \\ &= \sum_j w_{ji} \frac{\partial y_k}{\partial a_j} \end{aligned} \quad (6.15)$$

となる。微分 $\partial y_k / \partial a_j$ は再び連鎖律を用いることで

$$\begin{aligned} \frac{\partial y_k}{\partial a_j} &= \sum_l \frac{\partial y_k}{\partial a_l} \frac{\partial a_l}{\partial a_j} \\ &= h'(a_j) \sum_l w_{lj} \frac{\partial y_k}{\partial a_l} \end{aligned} \quad (6.16)$$

と表される。これは誤差逆伝播と同様な数式であり微分 $\partial y_k / \partial a_j$ が誤差の代わりにネットワークを伝播していくことになる。したがってヤコビ行列の各要素の計算も誤差逆伝播法により $O(W)$ の計算量で行うことができる。

6.2.2 hessian-vector-product

ニューラルネットワークのヘッセ行列 \mathbf{H} は重みベクトル \mathbf{w} に対する出力の二階微分として定義される。ヘッセ行列は誤差関数曲面の勾配の二次の情報まで使用する最適化アルゴリズムやネットワークの不要な重みを特定するために使われる。ヘッセ行列を計算するうえで大きな問題となるのがその次元の大きさである。ニューラルネットワークの重みパラメータの個数を W とするとヘッセ行列の要素数は W^2 となる。したがってネットワークが多数のパラメータを含む場合、ヘッセ行列の計算コストは非常に大きくなり、計算を実行できなくなる。そこでヘッセ行列の計算量を誤差逆伝播法により減らす研究がなされてきた。単純なものとしてはヘッセ行列の対角項のみを計算する対角近似やヘッセ行列をベクトルの外積の和で構成するレーベンバーグマーカード近似などがある。³⁸ いずれの手法も高次元のヘッセ行列を扱うために何かしらの近似や仮定を施しているが、もし興味のある対象が完全なヘッセ行列ではなく、ヘッセ行列 \mathbf{H} と何らかのベクトルの積 \mathbf{v} である場合、このヘッセ行列とベクトルの積(hessian vector product)は誤差逆伝播法を応用することで、近似を含まず、かつ $\mathcal{O}(W)$ の計算量で評価することができる。以下このアルゴリズムについて見ていく。

まずヘッセ行列 \mathbf{H} とベクトル \mathbf{v} の積 $\mathbf{v}^T \mathbf{H}$ を以下のように表す。

$$\mathbf{v}^T \mathbf{H} = \mathbf{v}^T \nabla(\nabla \mathbf{E}) \quad (6.17)$$

これは誤差関数の微分 $\nabla \mathbf{E}$ に演算子 $\mathbf{v}^T \nabla$ を作用させたものであると解釈することができる。この演算子 $\mathbf{v}^T \nabla$ を

$$\mathcal{R}\{\cdot\} = \mathbf{v}^T \nabla\{\cdot\} \quad (6.18)$$

と表すと、 $\mathbf{v}^T \mathbf{H}$ は $\nabla \mathbf{E}$ を誤差逆伝播により求める演算に、この $\mathcal{R}\{\cdot\}$ を作用させればよいことが分かる。まず順伝播方程式

$$\begin{aligned} a_j &= \sum_i w_{ji} x_i \\ z_j &= h(a_j) \\ y_k &= \sum_j w_{kj} z_j \end{aligned} \quad (6.19)$$

に $\mathcal{R}\{\cdot\}$ を作用させると、

$$\mathcal{R}\{\mathbf{w}\} = \mathbf{v} \quad (6.20)$$

であることを使いうと、

$$\begin{aligned} \mathcal{R}\{a_j\} &= \sum_i v_{ij} x_i \\ \mathcal{R}\{z_j\} &= h'(a_j) \mathcal{R}\{a_j\} \\ \mathcal{R}\{y_k\} &= \sum_j w_{kj} \mathcal{R}\{z_j\} + \sum_j v_{kj} z_j \end{aligned} \quad (6.21)$$

となることがわかる。ここで v_{ji} は \mathbf{v} の重みパラメータ w_{ij} に対する要素である。通常の順伝播方程式に含まれる変数に新たに、 $\mathcal{R}\{z_j\}$, $\mathcal{R}\{a_j\}$, $\mathcal{R}\{y_k\}$ が加わる。次に、逆伝播方程式に $\mathcal{R}\{\cdot\}$ を作用させる。誤差関数として二乗和誤差を

用いる場合,誤差 δ_k, δ_j は

$$\begin{aligned}\delta_k &= y_k - t_k \\ \delta_j &= h'(a_j) \sum_k w_{kj} \delta_k\end{aligned}\tag{6.22}$$

となるが,これに $\mathcal{R}\{\cdot\}$ を作用させると,

$$\begin{aligned}\mathcal{R}\{\delta_k\} &= \mathcal{R}\{y_k\} \\ \mathcal{R}\{\delta_j\} &= h''(a_j) \mathcal{R}\{a_j\} \sum_k w_{kj} \delta_k + h'(a_j) \sum_k v_{kj} \delta_k + h'(a_j) \sum_k w_{kj} \mathcal{R}\{\delta_k\}\end{aligned}\tag{6.23}$$

となる.最後に誤差関数の微分

$$\begin{aligned}\frac{\partial E}{\partial w_{kj}} &= \delta_k z_j \\ \frac{\partial E}{\partial w_{ji}} &= \delta_j z_i\end{aligned}\tag{6.24}$$

に $\mathcal{R}\{\cdot\}$ を作用させると

$$\begin{aligned}\mathcal{R}\left\{\frac{\partial E}{\partial w_{kj}}\right\} &= \mathcal{R}\{\delta_k\} z_j + \delta_k \mathcal{R}\{z_j\} \\ \mathcal{R}\left\{\frac{\partial E}{\partial w_{ji}}\right\} &= x_i \mathcal{R}\{\delta_j\}\end{aligned}\tag{6.25}$$

となり,hessian vector product $\mathbf{v}^T \mathbf{H}$ を得ることができる. このアルゴリズムを使用するには通常の順伝播,逆伝播方程式に $\mathcal{R}\{a_j\}, \mathcal{R}\{z_j\}, \mathcal{R}\{\delta_j\}, \mathcal{R}\{\delta_k\}, \mathcal{R}\{y_k\}$ を新たに変数として加えるだけでよく,容易に実装することができる.

6.3 CNNの物性計算への応用

畳み込みニューラルネットワーク (CNN) は画像認識分野における非常に強力なモデルであり,現在の人工知能ブームの火付け役となった. このCNNの大きなブームは,2012年の画像認識コンペティション ILSVRC において,トロント大学の Hinton らが開発した AlexNet が,従来の誤認識率を 10% 以上改善させたことに端を発する. 従来の画像認識では HOG や SIFT などの人の経験に基づく特徴量を使っていたが,CNN は学習データから自動的に特徴量を学習し抽出することで,高い性能を出すことに成功した. CNNは画像認識分野において誕生したものであるが,現在は画像にとどまらず,動画や音声,言語などにも応用され大きな成果を上げている. 物性予測においてもこのCNNを応用したモデルが提案されている. 従来の物性予測に使われてきたCoulomb matrixやPRDFなどは人の知見に基づく物質の特徴量であった. これらの特徴量を用いても物性予測において一定の成果を上げているが,CNNベースのモデルを使用することで,さらなる予測性能の向上が期待される. 問題となるのはCNNは画像を取り扱う学習モデルであるため,結晶構造をそのままCNNの入力とすることはできない. そこでCNNのモデルを改良し,結晶構造を取り扱えるようにしたモデルが提案された. この節ではそのようなCNNベースの結晶解析モデルCGCNNとR3DVSについて紹介する,^{39,40}

6.3.1 CGCNN

Xieらが提案したCrystal Graph Convolutional Neural Network(CG-CNN)の概要を図6.3に示す。

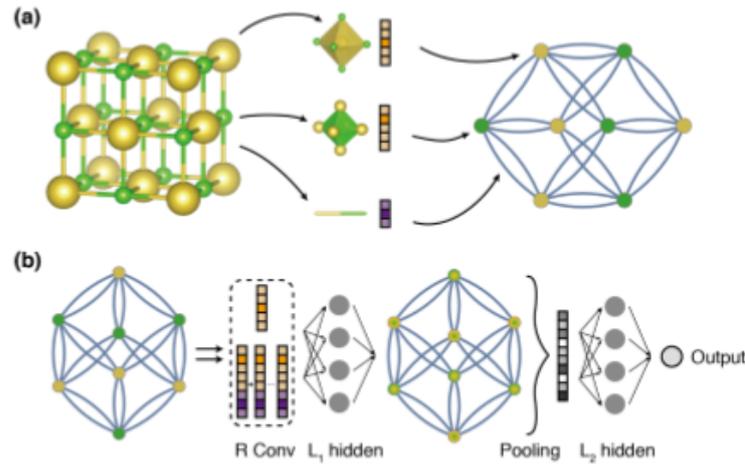


図6.3 CGCNNの構造

CNNに結晶構造の情報を入力するために,初めに結晶をグラフ構造に変換する. 各ノードは原子の特徴量ベクトル \mathbf{v}_i を有しており,各エッジは原子間の結合の情報 $\mathbf{u}_{(i,j)_k}$ を有している. このノードとエッジの特徴量ベクトルを使い,以下で定義する畳み込み操作を行うことで,新たな特徴量を生成する

$$\mathbf{v}_i^{(t+1)} = g \left[\left(\sum_{j,k} \mathbf{v}_j^{(t)} \oplus \mathbf{u}_{(i,j)_k} \right) \mathbf{w}_c^{(t)} + \mathbf{v}_i^{(t)} \mathbf{w}_s^{(t)} + \mathbf{b}^{(t)} \right]$$

この操作によって,結晶中の原子自身の情報とその周辺の原子の情報が統合された新たな特徴量ベクトルが生成される.

CNNは畳み込み操作に加え,入力データを粗視化するプーリング層を有している.CG-CNNにおけるプーリング操作は以下の式によって行われる.

$$\mathbf{v}_c = p \left(\mathbf{v}_0^{(0)}, \mathbf{v}_1^{(0)}, \dots, \mathbf{v}_N^{(0)}, \dots, \mathbf{v}_N^{(R)} \right)$$

上式の p はプーリング操作を表しており,論文中ではプーリング操作として入力となる特徴量の正規化した総和を用いている.

畳み込み層,プーリング層に加え, L_1, L_2 層の全結合ネットワークを用い,特徴ベクトルのマッピングを行い,出力となる物性値を得る. 予測値と教師データを使い,誤差逆伝播法により $\mathbf{W}_c^{(t)}, \mathbf{W}_s^{(t)}, \mathbf{b}^{(t)}$ を更新することで学習を行う.

CG-CNNの性能評価のため46744種類の無機結晶に対して結晶のエネルギー値やバンドギャップなどの物性値の予測を行った結果を図6.4に示す. 図6.4から分かるように,平均絶対誤差(MAE)の値は密度汎関数法による計算によって求めたものと同程度のものであり,高い精度で物性値の予測ができていることが分かる.

Property	# train data	Unit	MAE _{model}	MAE _{DFT}
Formation energy	28046	eV/atom	0.112	0.136[14]
Absolute energy	28046	eV/atom	0.227	0.19[17]
Band gap	16458	eV	0.530	0.6[17]
Fermi energy	28046	eV	0.564	–
Bulk moduli	2041	GPa	16.4	10.6[18]
Shear moduli	2041	GPa	13.4	6.86[18]
Poisson ratio	2041	–	0.033	–

図6.4 CGCNNによる物性値の予測結果

6.3.2 R3DVS

KajitaらはCNNを3次元に拡張した機械学習モデル3DCNNを用いた物性予測モデルR3DVSを構築した。R3DVSの構造を図6.5に示す。

モデルの入力となる三次元データは価電子密度 $\rho(\mathbf{r})$ を三次元フーリエ変換し、

$$|\rho(\mathbf{g})| = \frac{1}{v_c} \left| \int_{v_c} \exp(-i\mathbf{r} \cdot \mathbf{g}) \rho(\mathbf{r}) d\mathbf{r} \right| \quad (6.26)$$

得られた価電子密度 $\rho(\mathbf{g})$ を適当な波数のカットオフ g_{cut} までの逆格子空間上でメッシュに区切ったものを使った。さらにこのデータを回転させたレプリカを複数作成し、それらもデータセットに加えることで、モデルに回転不変性を持たせることを狙った。このモデルを物質の同定、すなわち入力されたデータがどの物質であるかという分類問題に適用した結果は図6.6のようになり、回転させたレプリカによるデータの補強によりより高い分類精度が出ていることが分かる。

このモデルを使い第一原理計算で計算された物性値を予測する実験を行った。600種類の無機結晶データそれぞれから30種類の回転させたレプリカを作成し、18000個のデータを用いてモデルを訓練した。その結果を図6.7に示す。電子の古典的なクーロン相互作用エネルギーであるハートリーエネルギーの予測では他の手法より優れた結果を残していることが分かる。

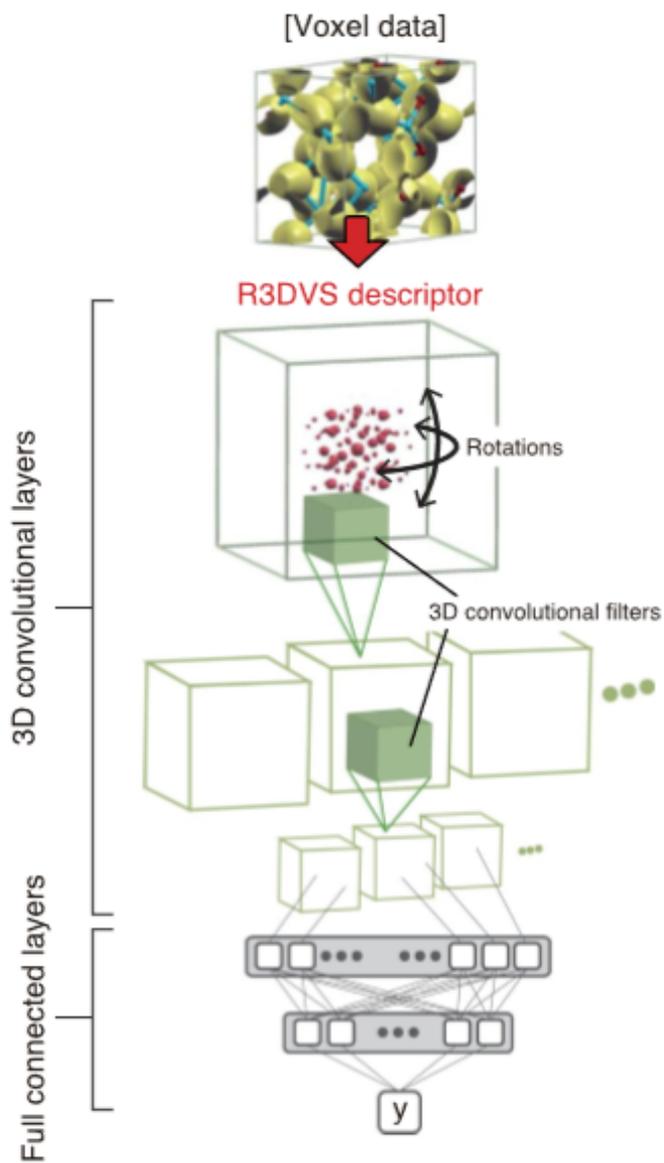


図6.5 R3DVSの模式図

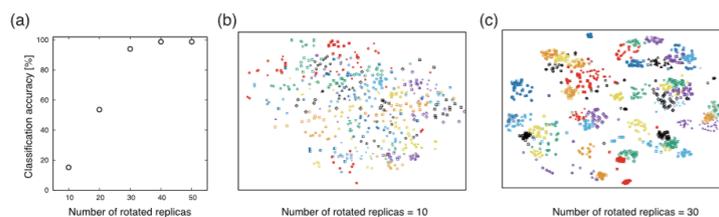


図6.6 R3DVSによって物質の同定実験を行った結果

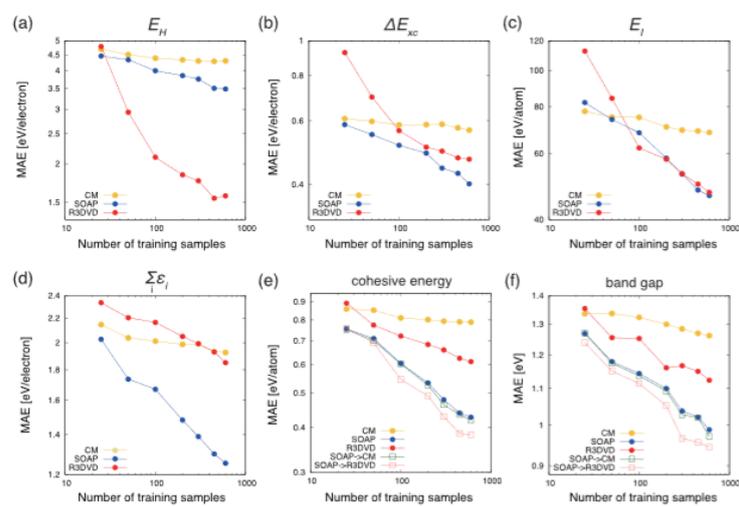


図6.7 R3DVSにより物性値を予測した結果

第7章

3DCNNによる交換相関汎関数法の開発

この章では3DCNNにより交換相関汎関数を開発するためのアルゴリズムについて述べる。まず用いたモデルの構造を示し、そのモデルによってどのようにして交換相関汎関数を構成するかについて述べ、このモデルの利点について説明する。次にモデルの学習アルゴリズムを導出する。最後に入力データとなる電子密度の扱いについて触れる。

7.1 3DCNNによる交換相関汎関数のモデリング

3DCNNにより交換相関汎関数を表現するには三次元の $n \times n \times n$ メッシュで切られた電子密度 $\rho(\mathbf{r})$ を入力とし、交換相関エネルギー E_{xc} を出力するネットワークを構成すればよい。交換相関ポテンシャル V_{xc} は

$$V_{xc} = \frac{\delta E_{xc}}{\delta \rho(\mathbf{r})} \quad (7.1)$$

と書けるが、これはネットワークのヤコビ行列を計算することで得られる。ヤコビ行列については先の章のアルゴリズムを使えば誤差逆伝播法により効率的に計算することができる。このネットワークを図7.1に示す。

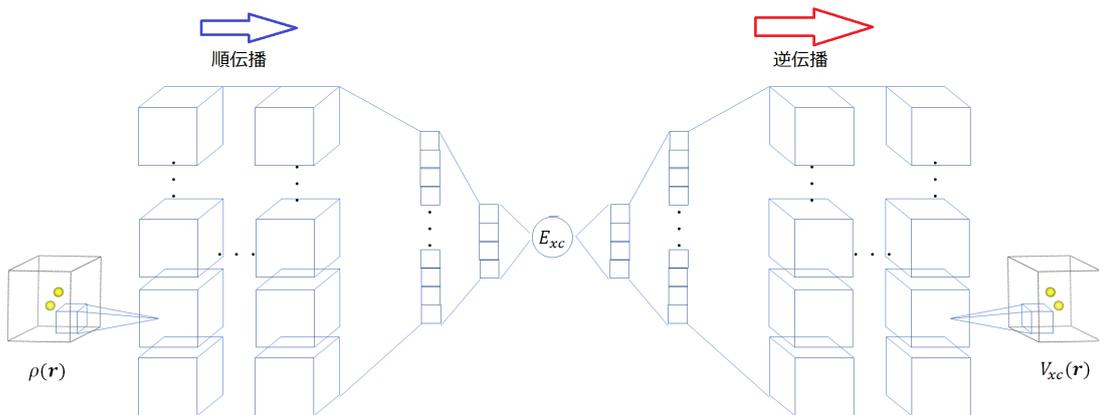


図7.1 3DCNNの構造

3DCNNにより交換相関汎関数を表現する最大の利点はその非局所性である。先の章で見てきたように交換相関汎関数は電子の非局所性を取り込もうと局所密度近似(ρ),一般化勾配近似($\nabla\rho$),metaGGA($\nabla^2\rho$)と拡張されて

いった。しかしmetaGGAであっても電子密度を ρ をメッシュで切った場合、隣の隣までの電子密度の情報しか取り入れることができない。一方で3DCNNの出力はすべての点からの情報が入っており、そのヤコビ行列を考えるとヤコビ行列の各要素は他のすべての点から影響を受ける。したがって、3DCNNで構成された交換相関ポテンシャルのある点の値は他のすべての点の電子密度を使って計算される。これはまさに完全に非局所的な関数であり、交換相関汎関数が追い求めている性質である。

また計算量においても大きな利点がある。ハートリーフォック法の交換相互作用項は非局所的な関数であるが、その計算を実行しようとするときある点の交換ポテンシャルを計算するのに他のすべての軌道との交換積分を実行せねばならず非常に計算コストが大きくなる。3DCNNの場合は電子密度 $\rho(\mathbf{r})$ を

$$\rho(\mathbf{r}) = \sum_i |\phi_i(\mathbf{r})|^2 \quad (7.2)$$

で計算し、それをネットワークに入力すれば $\mathcal{O}(W)$ 程度の計算量で計算することができる。電子密度の計算は交換積分を実行するよりもはるかに計算コストが低く、またネットワークの出力の計算も $\mathcal{O}(W)$ の単純な積和演算であるため高速に実行することができる。さらにネットワークの計算量はパラメータの数と同じである場合、電子密度のメッシュの細かさにしか依存しないため考えている物理系が単位格子であろうと大規模系であろうとそのスケールに依存することはない。密度汎関数計算の計算量は系に含まれる電子の総数を N とすると N^3 のオーダーであり、大規模系になると計算コストが爆発的に増える。近年は京などのスーパーコンピュータによって一万原子系の計算を行うことが可能になっているが、⁴¹密度汎関数計算の計算量を減らすことは理論の高精度化と並んで大きな課題である。もし既存の汎関数を3DCNNで完全にモデリングすることができれば、密度汎関数計算の計算コストを大幅に抑えることができると期待される。精度や物理的妥当性を保証することはできないが、密度汎関数計算における自己無撞着計算の初期値として3DCNNで計算されたものを与えれば、計算の収束を速めることができる可能性がある。

このように3DCNNで既存の交換相関汎関数をモデリングすることは、非局所性を少ない計算コストで取り入れることができ有用であるが、本研究では既存の交換相関汎関数をネットワークの重みパラメータを学習することで交換相関汎関数のさらなる高精度化を目指した。先の章でも論じたが、ニューラルネットワークは万能近似器 (universal approximation) であるため、密度汎関数理論の究極の目標ともいえる完全な交換相関汎関数を再現することが原理的には可能である。もちろん計算資源や精度による限界はあるが、ネットワークの深層化や電子密度データのメッシュの微細化などにより系統的に交換相関汎関数を改善していくことができる。現在の交換相関汎関数の開発においては、ヤコブの梯子を上ること以外の系統的な改良法はない。そのため、この3DCNNによる汎関数が成功を収めれば、究極の交換相関汎関数に近づくための新たなアプローチを与えることができる。この意味でも本研究の意義は大きい

本研究ではまず既存の交換相関汎関数を3DCNNでモデリングし,その重みパラメータを更新することでさらなる高精度化を試みた. 3DCNNにより交換相関汎関数をモデリングするためには,ネットワークの重みパラメータの最適化が必要である. 次節ではどのようにしてパラメータを更新していくかについて述べる.

7.2 ネットワークの訓練

まず,既存の交換相関汎関数を3DCNNでモデリングすることを考える. 一般的なニューラルネットワークの訓練は,ネットワークの出力と教師データによって構成される誤差関数の最小化によって行われる. 回帰問題では二乗和誤差(6.7)が使われることが多い. 本研究の場合ニューラルネットワークの出力は交換相関エネルギー $E_{xc}[\rho(\mathbf{r})]$ と交換相関ポテンシャル $V_{xc}(\mathbf{r})[\rho(\mathbf{r})]$ の二つである. 今回は既存の交換相関汎関数により計算されたそれぞれの値の二乗和誤差を計算し,それらの和を誤差関数とした.

$$E = \frac{1}{2} \{E_{xc}^{NN} - E_{xc}^{exact}\}^2 + \frac{1}{2} \int \{V_{xc}(\mathbf{r}) - V_{xc}^{exact}(\mathbf{r})\}^2 d\mathbf{r} \quad (7.3)$$

通常のニューラルネットワークの訓練を考えれば,誤差関数は $E_{xc}[\rho(\mathbf{r})]$ に関するもので十分であるように思われるが,交換相関汎関数で重要であるのはコーンシャム方程式の平均場を構成する交換相関ポテンシャル,すなわち $E_{xc}[\rho(\mathbf{r})]$ の入力 $\rho(\mathbf{r})$ に対する微分値であり,ネットワークの出力とその入力に対する微分が両方とも教師データを満足するものでないと正しい汎関数をモデリングしているとは言えない. 入力されるデータをデータ空間上で極めて稠密にとれば微分値に関しても正しく学習することができるが,モデルの選択や学習によっては大きな誤りを生む可能性がある. 下の例(7.2)の $M = 9$ の場合がこのことの良い例であり,過学習が起きたモデルでは出力は正確であるが,その微分値はデータを生成した関数のものとは大きく異なっている. 今回のように出力とその微分の情報を誤差関数に用いることでデータを生成した関数を精度よくモデリングすることが可能になる.

既存の汎関数のモデリングは上の議論でよいが,問題となるのはその先の汎関数の改良である. 最大の問題は交換相関エネルギーとポテンシャルは物理的な対応を持たない量であり実験値が存在しないため教師データとしてネットワークの訓練に使うことができないということである. したがってそれら以外の教師データとなる観測可能な物理量を探さなければならないが,その教師データとなる物理量のネットワークの重みパラメータによる微分値が計算できなければ,3DCNNの学習を行うことができない. しかしながら観測可能であり,かつネットワークのパラメータに関する微分値を計算できる物理量が存在する. それはバンドギャップである. バンドギャップは光学測定などで容易に測ることができる物理量であり,実験データも多い. もしバンドギャップを教師データとして用いることができれば,密度汎関数法の最大の欠点を克服することができる可能性がある. 本研究ではバンドギャップの値のネットワークのパラメータに関する微分値を計算する式を開発した. 以下ではその式の導出について述べる.

先の章でも述べたが密度汎関数法におけるバンドギャップはコーンシャム方程式を解いて得られるエネルギー

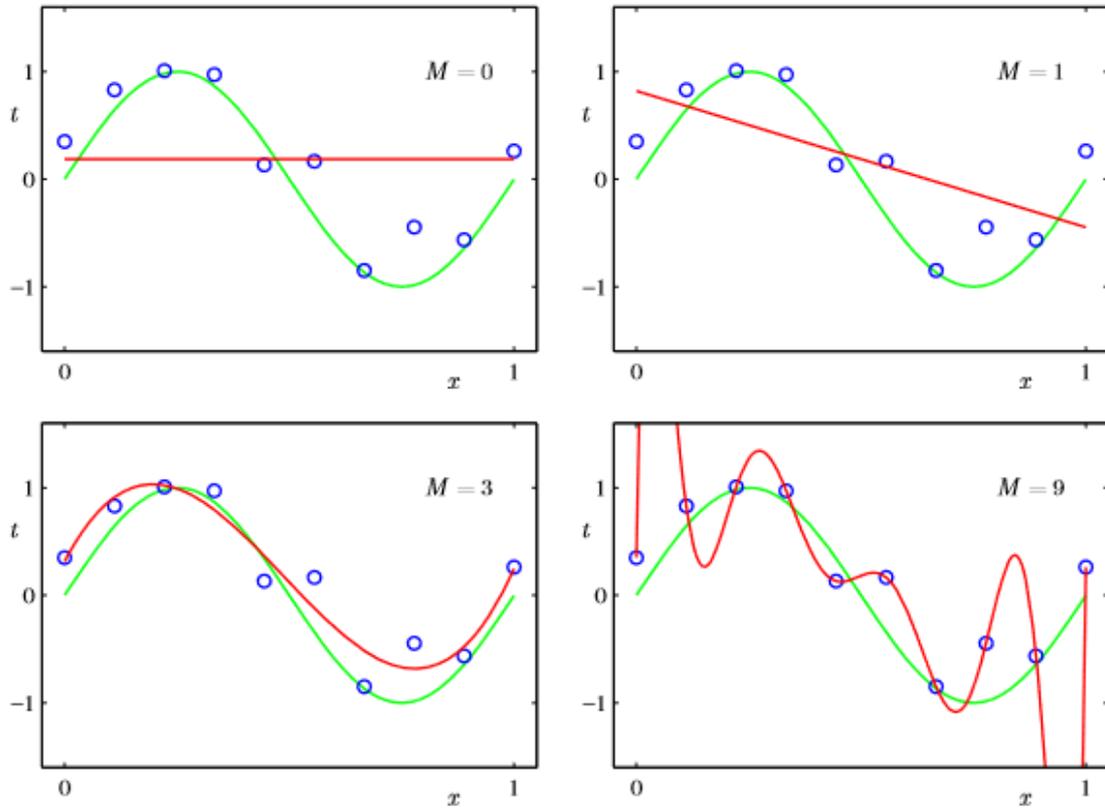


図7.2 過学習の例. $M = 9$ のモデルでは過学習が発生しており, 予想曲線が激しく振動している.³⁸

固有値の最低非占有準位 ε_{N+1} と最高占有準位 ε_N の差として定義される.

$$E_g = \varepsilon_{N+1} - \varepsilon_N \tag{7.4}$$

バンドギャップを教師データとして用いた場合, その誤差関数は

$$E = \frac{1}{2} (E_g - E_g^{\text{exp}})^2 \tag{7.5}$$

となる. その重みパラメータに関する微分は

$$\begin{aligned} \frac{\partial E}{\partial \mathbf{w}} &= (E_g - E_g^{\text{exp}}) \frac{\partial E_g}{\partial \mathbf{w}} \\ &= (E_g - E_g^{\text{exp}}) \left(\frac{\partial \varepsilon_{N+1}}{\partial \mathbf{w}} - \frac{\partial \varepsilon_N}{\partial \mathbf{w}} \right) \end{aligned} \tag{7.6}$$

となる. この式を見て分かるように重みパラメータに関する微分を計算するためにはコーンシャム準位の重みパラメータに関する微分が分かればよい. ここでコーンシャム方程式を重みパラメータ \mathbf{w} に関する依存性を分かるように書き下すと

$$\begin{aligned} \mathcal{H}(\mathbf{w})\phi_i(\mathbf{r}) &= \varepsilon_i\phi_i(\mathbf{r}) \\ \mathcal{H}(\mathbf{w}) &= -\frac{\hbar^2}{2m}\nabla^2 + V_{\text{eff}}(\mathbf{r}, \mathbf{w}) \\ V_{\text{eff}}(\mathbf{r}, \mathbf{w}) &= V(\mathbf{r}) + e^2 \int \frac{\rho(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r}' + V_{\text{xc}}(\mathbf{r}, \mathbf{w}) \end{aligned} \tag{7.7}$$

となる. したがってハミルトニアン $\mathcal{H}(\mathbf{w})$ は交換相関ポテンシャル $V_{xc}(\mathbf{r}, \mathbf{w})$ によって \mathbf{w} に関してパラメタライズされている. したがってこの問題はハミルトニアン $\mathcal{H}(\mathbf{w})$ のパラメータ \mathbf{w} を変化させたときその固有値 ε_i がどのように変化するかという問題として捉え直すことができる. ここで表記の簡便さのために固有関数をケット記号 $|\phi\rangle$ で書く. コーンシャム方程式は

$$\mathcal{H}(\mathbf{w}) |\phi_i\rangle = \varepsilon_i |\phi_i\rangle \quad (7.8)$$

となる. このコーンシャム方程式に左から $\langle\phi|$ を作用させ, 固有関数の規格直交条件 $\langle\phi_i | \phi_j\rangle = \delta_{ij}$ を使うと,

$$\varepsilon_i = \langle\phi_i | \mathcal{H}(\mathbf{w}) | \phi_i\rangle \quad (7.9)$$

となる. 固有値の重みパラメータ \mathbf{w} に関する微分は

$$\begin{aligned} \frac{\partial \varepsilon_i}{\partial \mathbf{w}} &= \frac{\partial}{\partial \mathbf{w}} \langle\phi_i | \mathcal{H}(\mathbf{w}) | \phi_i\rangle \\ &= \frac{\partial \langle\phi_i |}{\partial \mathbf{w}} \mathcal{H}(\mathbf{w}) | \phi_i\rangle + \langle\phi_i | \mathcal{H}(\mathbf{w}) \frac{\partial | \phi_i\rangle}{\partial \mathbf{w}} + \langle\phi_i | \frac{\partial \mathcal{H}(\mathbf{w})}{\partial \mathbf{w}} | \phi_i\rangle \\ &= \varepsilon_i \left(\frac{\partial \langle\phi_i |}{\partial \mathbf{w}} | \phi_i\rangle + \langle\phi_i | \frac{\partial | \phi_i\rangle}{\partial \mathbf{w}} \right) + \langle\phi_i | \frac{\partial \mathcal{H}(\mathbf{w})}{\partial \mathbf{w}} | \phi_i\rangle \\ &= \varepsilon_i \frac{\partial}{\partial \mathbf{w}} \langle\phi_i | \phi_i\rangle + \langle\phi_i | \frac{\partial \mathcal{H}(\mathbf{w})}{\partial \mathbf{w}} | \phi_i\rangle \\ &= \langle\phi_i | \frac{\partial \mathcal{H}(\mathbf{w})}{\partial \mathbf{w}} | \phi_i\rangle \\ &= \langle\phi_i | \frac{\partial V_{xc}(\mathbf{w})}{\partial \mathbf{w}} | \phi_i\rangle \end{aligned} \quad (7.10)$$

として計算することができ, 目的であった固有値のパラメータに関する微分を得ることができた. ここでブラケット記号を書き直すと,

$$\begin{aligned} \frac{\partial \varepsilon_i}{\partial \mathbf{w}} &= \langle\phi_i | \frac{\partial V_{xc}(\mathbf{w})}{\partial \mathbf{w}} | \phi_i\rangle \\ &= \int \frac{\partial V_{xc}(\mathbf{w})}{\partial \mathbf{w}} |\phi_i(\mathbf{r})|^2 d\mathbf{r} \end{aligned} \quad (7.11)$$

という積分の形になる. これと(7.6)を使うと誤差関数 E の重みパラメータ \mathbf{w} に関する微分は

$$\frac{\partial E}{\partial \mathbf{w}} = (E_g - E_g^{\text{exp}}) \int \frac{\partial V_{xc}(\mathbf{w})}{\partial \mathbf{w}} (|\phi_{N+1}(\mathbf{r})|^2 - |\phi_N(\mathbf{r})|^2) d\mathbf{r} \quad (7.12)$$

として得られ, 勾配降下法によりパラメータの値 \mathbf{w} を更新することができる. こうしてバンドギャップの値を改善するようにネットワークの学習を行うことができ, 既存の交換相関汎関数を超えた汎関数を得られる可能性がある. パラメータ更新式を得ることができたが, この式の中で問題となり得る箇所がある. それは交換相関汎関数 V_{xc} のパラメータ \mathbf{w} に関する微分の次元である. 交換相関汎関数 V_{xc} はネットワークの出力である交換相関エネルギー E_{xc} の微分であり, 入力データである電子密度 $\rho(\mathbf{r})$ の次元を持つ. 電子密度 $\rho(\mathbf{r})$ を $M \times M \times M$ のメッシュで切ったとすると, $V_{xc}(\mathbf{r})$ は M^3 の次元を持つ. したがって V_{xc} のパラメータ \mathbf{w} に関する微分はパラメータ数を N とすると, NM^3 の次元を持ち, 計算するには M^3 回の誤差逆伝播を行うことになる. 本研究ではメッシュ数として $M = 32$ としたが, この場合一つのデータに対して約三万回もの誤差逆伝播計算を行うことになり計算コストは非常に大きくなる. しかしこの問題は先の章で論じたhessian-vector-productを使うことで解決することができる. V_{xc} のパラメータ \mathbf{w} に関

する微分はヘッセ行列ではないが,同様の議論によりvector productの式を導くことができる. ここで(7.11)を見てみると, V_{xc} のパラメータ w に関する微分はギャップに関わる軌道の電子密度との積分の形になっている. 実際に数値計算を行う際は積分は二つのデータ同士の内積をとることで計算されるため,軌道の電子密度をhessianにかけられるベクトル, V_{xc} のパラメータ w に関する微分をhessianとみなせばhessian vector productのアルゴリズムにより $\mathcal{O}(N)$ の計算量で更新式を計算することができる.

以上の議論によりネットワークを学習するアルゴリズムが完成した.

7.3 入力データ

交換相関汎関数は電子密度 $\rho(\mathbf{r})$ を入力として交換相関エネルギー E_{xc} を出力する関数である. ここで入力データである電子密度 $\rho(\mathbf{r})$ について考えると,ネットワークの入力として非常に望ましくない特性を持っていることがわかる. すなわち電子密度は原子核近傍で特異点を持つ. 原子核近傍領域は非常にクーロンポテンシャルが深く,多数の電子が集まり内殻を形成している. そのため電子密度は図(7.3)のように内殻領域で急峻な変化しており,格子間領域と同等の扱いをすると数値的に著しい不安定性を生む. この領域を取り扱うために密度汎関数法では内殻領域のポテンシャルを近似したり,波動関数を平面波でなく球面調和関数で展開するなどの手法が存在する. 本研究では3DCNNにこの電子密度のデータを入力するが内殻領域をそのまま入力するとネットワークの出力誤差は極めて大きくなる. そこでこの内殻領域を完全に無視し,格子間領域の伝導電子のみを電子密度の入力データとして取り扱う. これにより系の大部分の電子を無視することになってしまうが,ネットワークの学習に大きな問題は生じないと考えられる. 電子の伝導はフェルミ面付近の電子に支配されており,内殻の電子は寄与しない. したがってネットワークの教師データとして使われるバンドギャップはフェルミ面近傍の伝導電子の振る舞いを知ることができれば十分精度よく再現できると考えられる. しかしながら,自己相互作用誤差は内殻電子の振る舞いを正しく取り扱うことで改善されるという報告もあり,⁵完全な交換相関汎関数を目指すには内殻電子を数値計算的に安定して取り扱う手法を開発する必要がある.

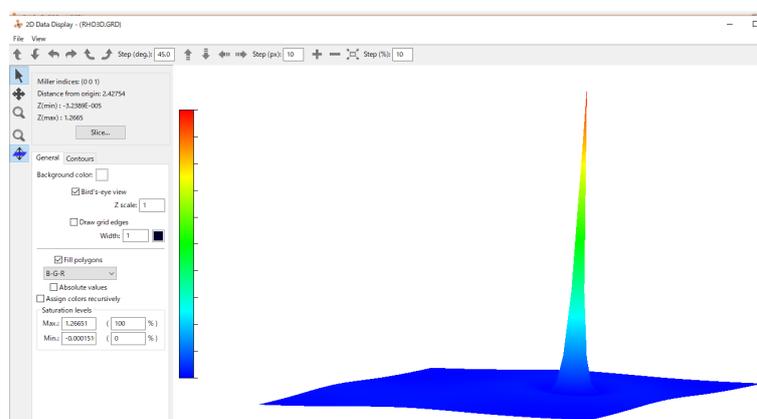


図7.3 原子核近傍の電子密度の変化

第8章

実験

先の章で3DCNNによる交換相関汎関数の開発アルゴリズムを見てきた。この章ではそのアルゴリズムに従い、実際にバンドギャップ誤差の補正実験を行った入力された結晶データの電子密度と交換相関エネルギー、ポテンシャルを計算に密度汎関数計算ライブラリElkを用いた。⁴² Elkは計算手法としてフルポテンシャル線形化APW法を用いており、第一原理計算ライブラリの中でも高い計算精度を誇る。⁴³ しかし計算時間は他の手法より多くなる傾向がある。

8.1 既存の汎関数のモデリング

まず初めに3DCNNを使った交換相関汎関数を開発するために、その雛形として既存の交換相関汎関数をモデリングする実験を行った。今回は固体物性計算で最もよく使われるPBE汎関数を対象として、そのモデリングを行った。

データセット

第一原理計算の入力となる結晶格子データを収集するために物質材料の第一原理計算結果のデータベースであるMaterials Projectを使用した。⁴⁴ Materials ProjectはMITの研究グループによって運営されており、結晶構造、バンド図、磁性などの第一原理計算の計算結果が多数収録されている。今回はMaterials ProjectのAPIを使い、単位格子に含まれる原子数が2以下の系の結晶構造データ(cifデータ)を1244個収集した。このcifデータを密度汎関数ライブラリの入力とし、PBE交換相関汎関数により伝導電子の電子密度とそれに対応する交換相関エネルギー、ポテンシャルを計算した。この電子密度と交換相関エネルギー、ポテンシャルのデータベースが3DCNNの入力と教師データとなる。

密度汎関数計算の設定

大量のデータを実行可能な時間内で計算できるように計算に用いる k 点メッシュは $1 \times 1 \times 1$ とした。 k 点のメッシュ数は細かいほど高精度な計算を行うことができるが、今回は単に電子密度と交換相関エネルギーの対応が得られれば良いので、計算時間を短縮できるよう最低限の k 点メッシュによって密度汎関数計算を行った。電子密度、交換相関ポテンシャルのメッシュは $32 \times 32 \times 32$ とした。

3DCNNの構造は先の章で紹介したR3DVSを参考に決定した.

layer	Filter size	Number of filters
Convolution	$3 \times 3 \times 3$	32
Relu	-	-
Average pooling	$2 \times 2 \times 2$	-
Convolution	$3 \times 3 \times 3$	32
Relu	-	-
Average pooling	$2 \times 2 \times 2$	-
Convolution	$3 \times 3 \times 3$	32
Relu	-	-
Average pooling	$2 \times 2 \times 2$	-
Full connected	8192	-
Full connected	4096	-
Full connected	1	-

図8.1 ネットワークの構造

このモデルを使用し,(7.3)で計算される誤差関数を最小化するようにモデルを訓練した. パラメータの最適化はAdam Optimizerにより1000回行った. 訓練データをを全体の七割, 検証データを一割, テストデータを全体の二割に分割した.

実験の結果, 図8.2のような学習曲線が得られた.

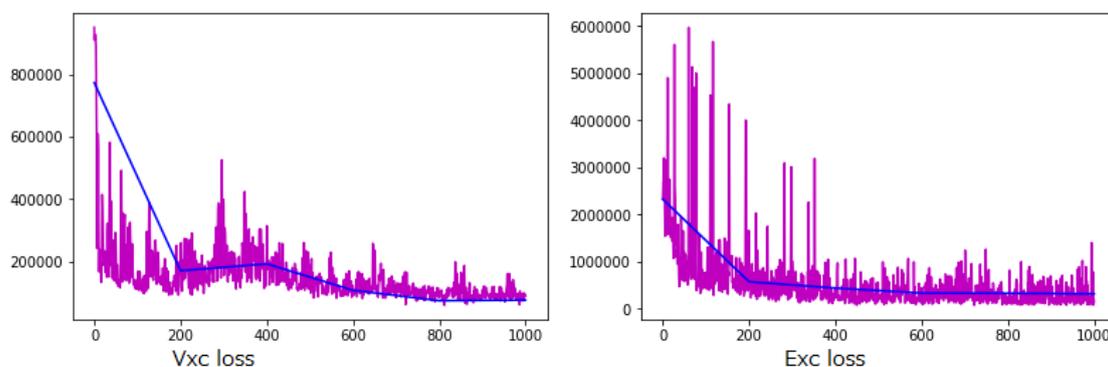


図8.2 3DCNNの学習結果

最終的に訓練データでは平均的な交換相関エネルギーの誤差は2eV程度, 交換相関ポテンシャルの各メッシュ点での誤差は1eV程度となった. これは化学計算を行う上では無視することができない程大きい誤差であるが, 本研究は既存の交換相関汎関数を雛形として新たな汎関数を開発することが目的であるため, この既存の汎関数とのずれはそれほど大きな問題ではない. こうして得られたPBE汎関数をモデリングした3DCNNによる交換相関汎関数を使い, バンドギャップの補正実験を行った.

8.2 既存汎関数でのバンドギャップの計算

先の章でも述べたように密度汎関数計算は自己無撞着計算を行う。もし自己無撞着計算に良い初期値が与えられれば、計算を速く収束させることができる。そこで先の実験で得られた3DCNNによる交換相関汎関数によるバンドギャップの補正実験を行う前に、既存の汎関数である程度正確な電子密度の計算を行っておく。3DCNN汎関数は既存の汎関数を模しているため、モデリングに使った汎関数によって計算された電子密度を初期値として使えば、一から自己無撞着計算を行うより速く計算を行うことができると期待される。今回は3DCNNのモデリングに使用したPBE交換相関汎関数を使用し、バンドギャップのデータセットに対して密度汎関数計算を行った。

データセット

バンドギャップのデータはcitrine informatics社から提供されている物性材料オープンデータベースcitrationから収集した。⁴⁵ citrine informatics社は材料開発を人工知能により加速することを目的としており、そのデータベースであるcitrationにはバンドギャップや熱力学特性などの様々な物性値のデータが収録されている。実験ではcitrationから1540件のバンドギャップのデータを収集した。収集したデータには同一の物質について複数の実験値を含んでいる。重複しているデータをまとめると、バンドギャップの実験値を得られた物質の種類は420件になった。その420件のデータに対しMaterials Projectから最安定状態の結晶構造データを収集し、密度汎関数計算を行った。単位胞中に多くの原子を含むものや、構造の対称性の低いものは計算時間がかかりすぎるため、一定の計算時間（120秒）が経過しても計算が終わらなかった物質はデータから除外した。このようにして最終的に209件のバンドギャップとその基底状態のデータセットが得られた。

密度汎関数計算の設定

バンドギャップを正確に推定するためには波数空間で細かいメッシュをとる必要があるが、細かい k 点メッシュをとるほど計算時間はそれに比例して大きくなる。通常の第一原理計算では k 点メッシュに対するエネルギーやバンドギャップの値の収束性から k 点メッシュの細かさを決定するが、今回のバンドギャップのデータベースに含まれる物質それぞれに対して k 点メッシュの収束性を考慮すると極めて時間がかかるため、今回はすべての物質に対し k 点メッシュを $4 \times 4 \times 4$ とした。物質の格子定数が小さくなるほど逆格子空間は大きくなり、精度良く計算を行うためには十分細かいメッシュをとらなければならない。今回は比較的小さな格子定数を持つダイヤモンド構造の炭素について密度汎関数計算を行い k 点のメッシュ数を決定した。これはデータベース中のすべての物質に対して計算精度的に十分細かいメッシュであるとは言えないが、平均的な格子定数を持つ物質に対しては十分であると考えられる。

上記の設定でPBE交換相関汎関数を使い密度汎関数計算を行った結果が図8.3である。

平均的な実験値とのバンドギャップの誤差は約1.17eVであった。密度汎関数法の欠点であるバンドギャップの

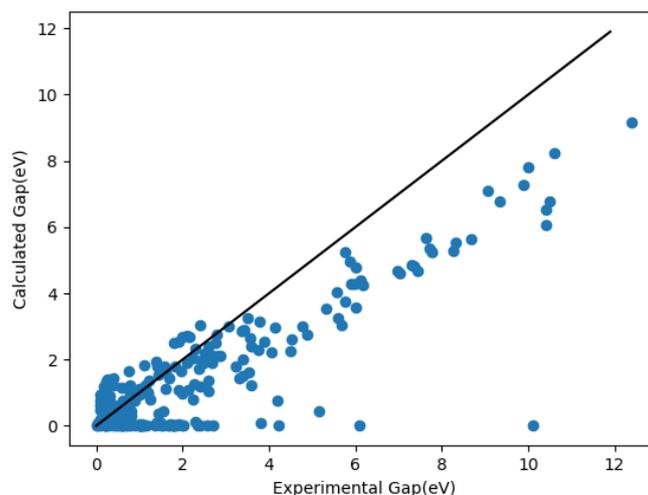


図8.3 PBEによるバンドギャップの計算結果

過小評価が見て取れる。またギャップの大きさが0でない物質に対してバンドギャップが消滅してしまっているものが見られる。これらは先の章のDFT+Uで述べたように電子相関により絶縁体化が引き起こされた物質であると考えられる。実際、NiOなどのバンドギャップを実験値が約4eVであるのに対し計算結果は限りなく0に近い値となっているが、NiOは典型的な強相関物質である。

8.3 小規模データセットでの検証

今回開発された学習アルゴリズムの正当性を検証するために、小規模なデータセットに対してギャップの補正実験を行った。用いたデータは塩化ナトリウム(NaCl)、塩化リチウム(LiCl)、フッ化リチウム(LiF)の3種類である。これらの物質はバンドギャップが非常に大きく、かつアルカリ金属のs軌道より構成されるきれいな放物上のバンドを伝導体に持つため、ギャップの補正検証としては結果の成否が分かりやすいと考えられ選ばれた。これらの物質に対し3DCNNを用いた汎関数を訓練するが、交換相関汎関数を完全に3DCNNのものに置き換えると数値的に不安定になる可能性があり計算が収束しなくなる恐れがあったため、今回は既存のPBE交換相関汎関数に1%だけ3DCNNによって計算される値を混合した。すなわち

$$\begin{aligned} V_{xc}(\mathbf{r}) &= (1 - \alpha)V_{xc}^{PBE}(\mathbf{r}) + \alpha V_{xc}^{NN}(\mathbf{r}) \\ E_{xc} &= (1 - \alpha)E_{xc}^{PBE} + \alpha E_{xc}^{NN} \end{aligned} \quad (8.1)$$

とし、 $\alpha = 0.01$ とした。この3DCNNで補正した新たな交換相関エネルギー、ポテンシャルを使い密度汎関数計算を行った。一回の訓練にすべてのデータを使うバッチ学習を200回行った結果図8.4のような学習曲線が得られた。

100ステップ目付近からギャップの誤差が大きく減少していることが分かる。また最終的に誤差が0付近に収束していく様子が見て取れ、この学習アルゴリズムが正しく働いていることが証明された。

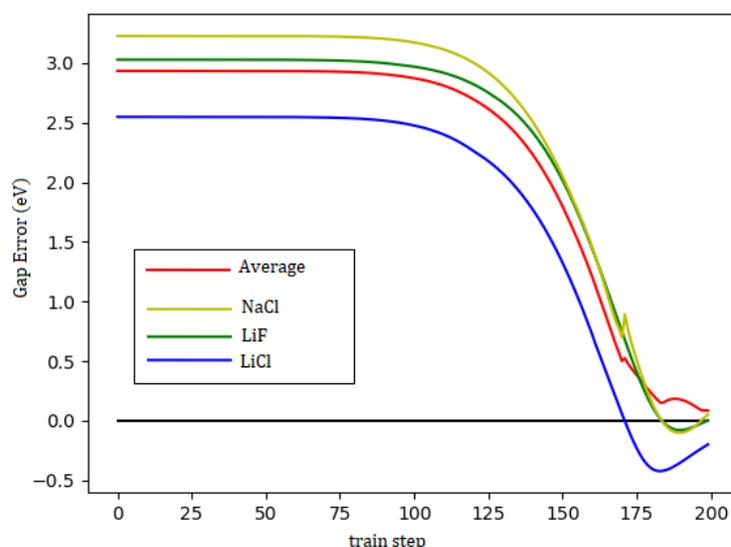


図8.4 NaCl LiCl LiFに対してバンドギャップを補正した結果

8.4 バンドギャップの補正

先の実験で補正アルゴリズムの正当性は実証されたため、3DCNN汎関数を使い実際のデータセットに対して補正実験を行った。先の実験より、ギャップが0に収束するまで約0.5eV程の振れ幅を持つことが分かった。もしギャップが0.5eV以下の物質であれば学習中にバンドギャップが潰れてしまう可能性がある。バンドギャップが潰れると、最低非占有軌道、最高占有軌道がそれまでギャップを形成してきたものから劇的に変化するため学習に悪影響が生じる可能性がある。今回は余裕をもって、PBE汎関数による予測値と実験値の値の両方が2eV以上である物質を対象に実験を行った。3DCNN汎関数の混合率を先の実験と同様に1%とし、バッチサイズを3、学習回数を300回とした。

実験の結果以下の図8.5のようになった

図を見て分かるように3DCNNによりバンドギャップ誤差を改善することはできず、むしろ悪化してしまっている。バンドギャップの平均的な実験値との誤差はPBE汎関数が約1.60eVであるのに対し、3DCNNの誤差は約2.06eVまで大きくなっている。

原因の一つとしては多数のデータセットに対して共通の補正汎関数を見つけ出すことができなかったことが考えられる。少数のデータセットに対してはうまくそれらに共通する補正汎関数を見つけ出すことができていたが、データが増えるとその共通項を見つけ出すことが難しくなり、バッチに対しては補正する方向に重みパラメータを更新できるがその更新された重みパラメータが他のデータの補正に効いてこなかったため、何度もパラメータの更新を重ねることでこのように全体的にずれていったと考えられる。この問題を解決するにはバッチサイズをより大きくすることが考えられる。バッチサイズを大きくすれば重みパラメータを更新する際、より多くのデータに共通

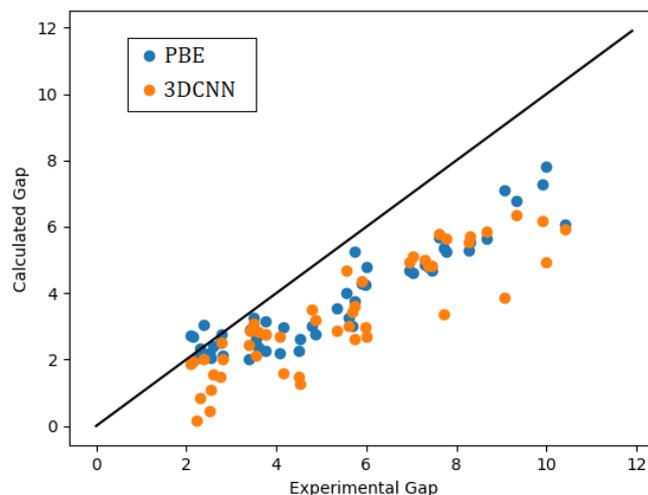


図8.5 データセットのバンドギャップを補正した結果

する改善方向に重みを更新することができるかと期待される。しかしバッチサイズを大きくすることの問題は重みパラメータを更新するのに計算時間がかかるということである。今回の補正アルゴリズムの計算時間の最大のボトルネックはバッチごとに密度汎関数計算をしなければならず、この計算は重みパラメータの更新に比べはるかに計算時間がかかる。重みパラメータの更新にはバンドギャップと最高占有軌道、最低非占有軌道の電子密度が必要であるが、これらを線形近似などで密度汎関数計算を行わずに求めることができれば、計算時間は大幅に短縮できる。今回はバンドギャップに関して重みパラメータに対する微分を求めることができたため、重み更新量と微分値によりバンドギャップに関しては線形近似を施すことができる。軌道の重みパラメータ依存性に対する微分が計算できればバンドギャップと軌道に関して線形近似によりパラメータ更新後の値を知ることができ、十分ゆっくりと学習を行えば密度汎関数計算を行う回数を減らすことができる可能性がある。軌道の重みパラメータ依存性の研究は今後の課題としたい。

補正実験がうまくいかなかったもう一つの原因としてギャップの大きさの異なる領域のデータをすべてまとめてデータセットに入れてしまったことが考えられる。少数データセットの実験を見て分かるように、学習が進んでいるときは一回の重みパラメータの更新で0.1eV程度ギャップの値が変化している。もしデータセットのギャップの小さい領域のデータが重みパラメータの更新により0.1eV程度変化すると、実際のギャップの大きさに対する変化の割合はギャップが大きいものに比べ大きい。このような変化が繰り返されるとギャップが小さい領域のデータの教師データへの収束は不安定になってしまう。またギャップが大きい領域のデータに対しても、小さい領域での誤差が大きいままであるため、Adam Optimizerが教師データへ高速に収束することを阻害されるため、いつまでも解となるデータへ収束することができなくなってしまう。この問題の解決法としては、データ領域を分けて学習を行うことが考えられる。ギャップの大きい領域のデータのみで十分に学習された汎関数を使い、ギャップの小さい

領域のデータに対して学習を行えば、ギャップの小さい領域のデータは誤差そのものも小さいため、ネットワークへの変化は微小であると考えられる。そのため再びギャップの大きいデータで学習を行ってもある程度は改善された状態であることが期待される。ギャップの大きい領域と小さい領域の学習を交互に繰り返せばデータセット全体のギャップを改善することができる可能性がある。問題となるのは一方の領域で過学習が起り、もう一方の領域に移った時に誤差が悪化する可能性があるということである。したがって一方の領域で過度に教師データにフィットさせすぎないように注意する必要がある。

本研究ではバンドギャップを補正することで既存の交換相関汎関数を越えた新たな汎関数を開発することを目指したが、結果としてそのような汎関数を得ることはできなかった。しかしながらギャップを補正するようにネットワークの重みパラメータを更新するアルゴリズムの正当性は保証されたため、今後データの扱いや学習の進め方を工夫することで、あらゆるバンドギャップを正しく再現する交換相関汎関数を開発することができる可能性がある。

第9章

結論

本研究では3DCNNにより既存の交換相関汎関数を超える新たな汎関数を開発することを目指した。汎関数を開発する指針としてまず初めに既存の汎関数のモデリングを行った。既存の汎関数のモデリングは化学計算的には十分な精度ではないものの、ある程度雛形となる汎関数へと3DCNNを収束させることに成功した。その得られた3DCNNモデルを使い、実際の物性値を再現するようにネットワークの重みパラメータを更新する実験を行った。再現すべき物性値としてバンドギャップを採用することで、パラメータの更新式を導くことに成功した。小規模なデータセットに対してこのパラメータ更新式によりモデルの学習を行ったところ、実際のバンドギャップの値へと収束していく様子が確認され、更新式の正当性を証明することができた。更新式の正当性が証明されたため、データセット全体に対して学習を行ったが結果は既存の汎関数を改善することはできず、むしろ悪化させてしまうことになった。このことはデータセット全体の誤差を改善する共通成分を見つけ出せなかったことが原因であると考えられるためバッチサイズの工夫などで対処する必要がある。またギャップの値の大きさにデータセットを分割し、分割されたデータセットにおいて転移学習をすることで、ギャップデータ全体の学習を速く安定して行うことができると期待される。

本研究の最終目標である既存の汎関数を超える交換相関汎関数を開発することはできなかったが、それを実現するための手法の正当性を確認することはできた。したがってデータの取り扱いや学習を工夫することで、今回の実験結果を改善することができる可能性がある。バンドギャップを高精度に再現することができる交換相関汎関数を開発することができたとしても、それが厳密な交換相関汎関数であるという保証は全くないが、あらゆる物性値を再現することができるという厳密な汎関数が満たすべき条件の一部を満足していることになるので、その候補となることはできる。今回開発した学習アルゴリズムはバンドギャップのみを教師データとしたが、バンドギャップ以外の物性値とネットワークの重みパラメータを関連付けることができれば、より多くの物性値を満足することができるモデルを構築することができ、厳密な汎関数に近づくことができる。また実験値だけでなく、高密度極限や低密度極限の電子密度に対する厳密解を教師データとして学習すれば、基礎物理条件をある程度満足するモデルを構築することができる。通常の交換相関汎関数では設計時にどの基礎物理条件を満たすかが決まっているので、汎関数中のパラメータをフィッティングしたとしても設計時に考慮されていない基礎物理条件を満たすことはできない。本

研究で開発されたモデルは学習によってどのような関数にも変化することができるので基礎物理条件を満たすようにしながらデータへのフィッティングを行うことが可能である。今回はモデルの学習において基礎物理条件に関して何の制約も加えなかったが、厳密な交換相関汎関数を開発するためにはそれを考慮しながら学習を行う必要があるため、今後の課題としたい。

今回開発したモデルは電子密度を入力として交換相関エネルギーを出力するものであったが、入力に磁化などの電子密度以外の物理量を加えれば交換相関汎関数を容易に拡張することができる。加えた基本変数に対して微分が定義できる物理量が存在すれば交換相関汎関数の学習を行うことができ、拡張された密度汎関数理論を構築することができる。以前の章で述べたように、密度汎関数法を拡張するには、拡張された基本変数に対する交換相関汎関数を開発しなければならなかった。今回の手法が成功を収めれば、拡張された密度汎関数法における交換相関汎関数の開発がも容易に行うことができるようになり、密度汎関数法をさらに発展させることができる。今回行った実験がうまくいった暁には、基本変数に電子密度以外の物理量を加えることでモデルを拡張していきたい。

厳密な交換相関汎関数は物性物理の究極の目標であり、もしそれが手に入ればあらゆる物質を第一原理的に厳密に解析することができる。本研究はその究極の目標に迫る新たなアプローチを与える研究であるといえる。今回はうまくいかなかったが、手法の改良によりいつか厳密な汎関数を手に入れることができる可能性がある。

謝辞

今回、研究を進めるにあたって多くの方々にご協力いただきました。この場を借りて感謝の意を表したいと思います。

指導教官である伊庭斉志教授には普段のミーティングにおいて研究に関して様々な助言を頂きました。また研究以外の人工知能や進化に関する話をたくさんして頂きました。物理工学出身の私がこの分野に興味を持って研究に取り組むことができたのは先生のような話が大変興味深く思えたからです。本当にありがとうございました。

助教の土肥浩先生にはハードウェアからソフトウェアの話まで様々なことを学ばせて頂きました。そのおかげでコンピュータに対する理解が深まり、円滑に研究を進めることができました。

博士課程1年のピンドゥル康太郎氏には物理に関する様々な相談に乗って頂きました。情報系の学科で物理の相談ができる方がいらっしやったことは大変ありがたかったです。

長谷川研究室博士課程2年の計良宥志氏には研究室に入ってから現在まで数多くの助言を頂きました。今私が研究できているのは氏のおかげといっても過言ではありません。

研究室の他のメンバーにも普段のミーティングや雑談で様々な助言を頂きました。

ここで紹介した方以外にも多くの方々に支えられて研究を進めることができました。ここに感謝申し上げます。

参考文献

- [1] Tansley, Stewart, and Kristin M. Tolle. "The fourth paradigm: data-intensive scientific discovery", Ed. Tony Hey. Vol. 1. Redmond, WA: Microsoft research, (2009).
- [2] 赤井久純,白井光雲, 密度汎関数法の発展-マテリアルデザインへの応用, 丸善出版(2012).
- [3] Löwdin, Per-Olov (March 1955). "Quantum Theory of Many-Particle Systems. III. Extension of the Hartree–Fock Scheme to Include Degenerate Systems and Correlation Effects". *Physical Review*. 97 (6): 1509–1520
- [4] 夏目雄平,小川建吾, 計算物理II, 朝倉書店(2002).
- [5] 常田貴夫, 密度汎関数法の基礎, 講談社サイエンティフィク(2012).
- [6] Møller, Chr, and Milton S. Plesset. "Note on an approximation treatment for many-electron systems." *Physical Review* 46.7 (1934): 618.
- [7] F. Sato, T. Yoshihiro, M. Era and H. Kashiwagi, "Calculation of all-electron wavefunction of hemoprotein cytochrome c by density functional theory," *Chem. Phys. Lett.*, 341(2001), pp. 645-651
- [8] Geng, Ming, and Hannes Jónsson. "Density functional theory calculations and thermodynamic analysis of bridgmanite surface structure." *Physical Chemistry Chemical Physics* (2019).
- [9] Umemoto, Koichiro, Renata M. Wentzcovitch, and Philip B. Allen. "Dissociation of MgSiO₃ in the cores of gas giants and terrestrial exoplanets." *Science* 311.5763 (2006): 983-986.
- [10] David S. Sholl, Janice A. Steckel, 佐々木泰造,末原茂 共訳 密度汎関数理論入門 理論とその応用, 吉岡書店(2014).
- [11] Hohenberg, Pierre, and Walter Kohn. "Inhomogeneous electron gas." *Physical review* 136.3B (1964): B864.
- [12] 押山淳,天能精一郎,杉野修,大野かおる,今田正俊,高田康民, 計算と物質, 岩波書店(2012).
- [13] Kohn, Walter, and Lu Jeu Sham. "Self-consistent equations including exchange and correlation effects." *Physical review* 140.4A (1965): A1133.
- [14] 小口多美夫, バンド理論：物質科学の基礎として , 内田老鶴圃(1999).
- [15] Schönhammer, K., O. Gunnarsson, and R. M. Noack. "Density-functional theory on a lattice: Comparison with exact numerical results for a model with strongly correlated electrons." *Physical Review B*

- 52.4 (1995): 2504.
- [16] von Barth, Ulf, and Lars Hedin. "A local exchange-correlation potential for the spin polarized case. i." *Journal of Physics C: Solid State Physics* 5.13 (1972): 1629.
- [17] Vignale, G., and Mark Rasolt. "Density-functional theory in strong magnetic fields." *Physical review letters* 59.20 (1987): 2360.
- [18] Oliveira, L. N., E. K. U. Gross, and W. Kohn. "Density-functional theory for superconductors." *Physical review letters* 60.23 (1988): 2430.
- [19] Carchini G. et al. . How theoretical simulations can address the structure and activity of nanoparticles. *Top. Catal.* 56, 1262–1272 (2013) .
- [20] Vosko, Seymour H., Leslie Wilk, and Marwan Nusair. "Accurate spin-dependent electron liquid correlation energies for local spin density calculations: a critical analysis." *Canadian Journal of physics* 58.8 (1980): 1200-1211.
- [21] Perdew, John P., and Yue Wang. "Accurate and simple analytic representation of the electron-gas correlation energy." *Physical Review B* 45.23 (1992): 13244.
- [22] Becke, Axel D. "Density-functional exchange-energy approximation with correct asymptotic behavior." *Physical review A* 38.6 (1988): 3098.
- [23] Perdew, John P., et al. "Atoms, molecules, solids, and surfaces: Applications of the generalized gradient approximation for exchange and correlation." *Physical Review B* 46.11 (1992): 6671.
- [24] Perdew, John P., Kieron Burke, and Matthias Ernzerhof. "Generalized gradient approximation made simple." *Physical review letters* 77.18 (1996): 3865.
- [25] Accurate Density Functional with Correct Formal Properties: A Step Beyond the Generalized Gradient Approximation
- [26] Stephens, P. J., et al. "Ab initio calculation of vibrational absorption and circular dichroism spectra using density functional force fields." *The Journal of Physical Chemistry* 98.45 (1994): 11623-11627.
- [27] Becke, Axel D. "Density-functional thermochemistry. V. Systematic optimization of exchange-correlation functionals." *The Journal of chemical physics* 107.20 (1997): 8554-8560.
- [28] Zhao, Yan, and Donald G. Truhlar. "A new local density functional for main-group thermochemistry, transition metal bonding, thermochemical kinetics, and noncovalent interactions." *The Journal of chemical physics* 125.19 (2006): 194101.
- [29] Zhao, Yan, and Donald G. Truhlar. "The M06 suite of density functionals for main group thermochemistry, thermochemical kinetics, noncovalent interactions, excited states, and transition elements: two new functionals and systematic testing of four M06-class functionals and 12 other functionals."

- Theoretical Chemistry Accounts 120.1-3 (2008): 215-241.
- [30] Ward, L. et al, "Matminer: An open source toolkit for materials data mining." *Comput. Mater. Sci.* 152, 60-69 (2018).
- [31] 高田康民, 多体問題特論 : 第一原理からの多電子問題, 朝倉書店(2009).
- [32] Adamo, Carlo, and Vincenzo Barone. "Toward reliable density functional methods without adjustable parameters: The PBE0 model." *The Journal of chemical physics* 110.13 (1999): 6158-6170.
- [33] Heyd, Jochen, Gustavo E. Scuseria, and Matthias Ernzerhof. "Hybrid functionals based on a screened Coulomb potential." *The Journal of chemical physics* 118.18 (2003): 8207-8215.
- [34] A. J. Garza and G. E. Scuseria, "Predicting band gaps with hybrid density functionals", *J. Phys. Chem. Lett.* 7, 4165-4170 (2016).
- [35] Hedin, Lars. "New method for calculating the one-particle Green's function with application to the electron-gas problem." *Physical Review* 139.3A (1965): A796.
- [36] M. van Schilfgaarde, T. Kotani, and S. Faleev, "Quasiparticle Self-Consistent GW Theory", *Phys. Rev. Lett.* 96 (2006). 226402.
- [37] Loschen, C. et al, "First-Principles LDA + U and GGA + U Study of Cerium Oxides: Dependence on the Effective U Parameter" *PhysRevB*.75.035115
- [38] C.M.ピシヨップ, 村田昇 監訳, パターン認識と機械学習 上, 丸善出版(2012).
- [39] Xie, Tian, and Jeffrey C. Grossman. "Crystal Graph Convolutional Neural Networks for Accurate and Interpretable Prediction of Material Properties." *arXiv preprint arXiv:1710.10324* (2017).
- [40] Seiji Kajita. et al, "A Universal 3D Voxel Descriptor for Solid-State Material Informatics with Deep Convolutional Neural Networks" *Scientific Reports*, 7, 16991 (2017)
- [41] 宮崎 剛, 百万原子系に対する第一原理分子動力学シミュレーションの実現 : オーダーN 法第一原理計算プログラム CONQUEST の最近の発展, *アンサンブル*, 18, 2, 74-81 (2016)
- [42] <http://elk.sourceforge.net/>
- [43] <https://molmod.ugent.be/deltacodesdft>
- [44] A. Jain. et al "The Materials Project: A materials genome approach to accelerating materials innovation", *APL Materials*, 2013, 1(1), 011002.
- [45] <https://citration.com>.

対外発表

堀場貴裕,伊庭斉志. "パラメトリック遺伝的プログラミング" 進化計算学会第11回進化計算シンポジウム2017講演論文集 pp.49-56(2017)