計算生命科学の基礎∨| 生命科学のためのシミュレーション技術とデータサイエンス: 基礎から医療・創薬・健康科学への応用まで

# 量子化学計算の現在と近未来

#### 常田貴夫

神戸大学科学技術イノベーション研究科

#### 2019年11月6日 @神戸大学計算科学教育センター

#### 科学における演繹と帰納



Georg Wilhelm Friedrich Hegel



Albert Einstein

#### 科学においては, 理論にもとづく演繹的な推論の 重要性が強調されてきた

「実験結果のみからの帰納的推論は論理的な 演繹なしに科学的事実と認めえない」(ヘーゲル) →オームの法則は数十年間,科学的事実として 認められなかった

「自然の理解における真に偉大な進歩は帰納的 手法とほとんど正反対のところにある」 (『物理学における帰納と演繹』, アインシュタイン)

「実験は、科学的事実を確かめる唯一の手段 である」(ファインマン)

理論による演繹的推論は,近年の科学において は特に日本において軽視されてきた

実験結果に矛盾しない理論計算からの考察は, 実験結果のみからの考察より優位にある



Georg Simon Ohm



**Richard Feynman** 

# 量子化学の貢献できる研究分野



# 科学と理論化学の主題

期間	科学の主題	理論化学	科学全体
1926 – 1937	量子力学の基礎理論	HF法, 経験法, VB法, 摂動法, 固体論, 反応論など	量子論, 相対論的理論, 原子核論, 核融合など
1938 – 1949	原子爆弾関連	反応速度論など	核分裂, 超ウラン元素, 超流動, 原爆など
1950 - 1960	コンピュータと量子力学	ローターン法, 基底関数, 古	大型計算機, レーザー, DNA,
	の応用	典MD法, 電子相関法など	超伝導理論など
1961 - 1968	具体的な対象のための	軌道論, 線形応答論, CC法,	タンパク質, カオス, クォーク,
	技術・理論	多配置法, GVB法など	RNAなどの解明
1969 - 1984	汎用コンピュータと生体	DFT, MDポテンシャル, 励起	パソコン, 遺伝子, 塩基配列
	機能の解明	状態理論, Gaussian公開など	解読, 固体電子論など
1985 - 1995	ナノマテリアルと光化学	第1原理MD法, DFT汎関数, 多参照法, MDアンサンブル	フェムト秒などレーザー科学, 高温超伝導, C60など
1996 – 2005	技術・理論の融合によ	Order-N化法, QM/MM法,	フェムト秒科学応用, アト秒
	るユーティリティの追求	DFT補正法など	科学, ES細胞, ゲノム科学など
2006- ?	不可能を可能にする	軌道エネルギー, 反応経路	iPS細胞, エピジェネティクスな
	技術・理論?	探索, AI予測モデルなど?	ど

## 量子化学計算の利用頻度

#### 90年代から利用されはじめた密度汎関数法(DFT)が,現在では量子化学計算の主流



# 多電子問題の解きかた:ハートリー法



ハートリー近似 独立電子近似にもとづく方法 波動関数  $\Psi = \phi_1(\mathbf{r}_1) \cdot \phi_2(\mathbf{r}_2)$ ハミルトニアン演算子  $\hat{H} = \hat{h}(\mathbf{r}_1) + \hat{h}(\mathbf{r}_2)$ 

<u>ハートリー方程式</u> 多電子からの寄与はすべてポテンシャルで近似する  $\hat{h}(\mathbf{r}_{1})\phi_{1}(\mathbf{r}_{1}) = \left\{-\frac{1}{2}\nabla_{1}^{2} - \frac{2}{r_{1A}} + V_{\text{eff}}(\mathbf{r}_{1})\right\}\phi_{1}(\mathbf{r}_{1}) = \varepsilon_{1}\phi_{1}(\mathbf{r}_{1})$  $\hat{h}(\mathbf{r}_{2})\phi_{1}(\mathbf{r}_{2}) = \left\{-\frac{1}{2}\nabla_{2}^{2} - \frac{2}{r_{2A}} + V_{\text{eff}}(\mathbf{r}_{2})\right\}\phi_{2}(\mathbf{r}_{2}) = \varepsilon_{2}\phi_{2}(\mathbf{r}_{2})$ 



D. R. Hartree

非線形方程式=自己無撞着場(SCF)法で解けばよい

## 3体問題



図. 3体運動のシミュレーション



3体の運動をシミュレーションすると、 すくなくとも1つが超高速ではじき出される

図. 平面上で逆回りしている2つのペアの間を 1つの小粒子が往復する質点モデル

2つのペアは無限の速さで遠ざかる

### 水素分子のなかの電子運動



#### 分子軌道法

LCAO-MO近似: 分子のなかの電子運動の経路=分子軌道を, 原子のなかの電子運動の経路=原子軌道の線形結合であらわす



# 分子軌道に対する変分原理

## パウリ排他原理を満たす波動関数



P. A. M. Dirac

# 反対称化波動関数にもとづく解きかた

	<u>ノーターイ</u>	丁列式		
$\Psi(\mathbf{r}_1, \mathbf{r}_2, \cdots, \mathbf{r}_N) = \frac{1}{\sqrt{(2n)!}}$	$ \begin{array}{c} \phi_1(\mathbf{r}_1)\alpha \\ \phi_1(\mathbf{r}_2)\alpha \\ \vdots \\ \phi_1(\mathbf{r}_{2n})\alpha \end{array} $	$ \begin{array}{c} \phi_1(\mathbf{r}_1)\beta \\ \phi_1(\mathbf{r}_2)\beta \\ \vdots \\ \phi_1(\mathbf{r}_{2n})\beta \end{array} $	···· ··· ··.	
$=\frac{1}{\sqrt{(2n)!}}$	$\det  \phi_1 \alpha \phi_1\rangle$	$\beta \cdots \phi_n eta$		



J. C. Slater

ハートリー・フォックエネルギー:スレーター行列式に対するエネルギー期待値

$$E = \frac{\int d\tau \Psi^* \hat{H} \Psi}{\int d\tau \Psi^* \Psi} = 2\sum_{i=1}^n h_i + \sum_{i,j=1}^n \left(2J_{ij} - K_{ij}\right)$$

多電子寄与をポテンシャルではなくあらわな軌道間相互作用とする=ab initio法

#### Ab initio ハートリー・フォック法

ハートリー法へのスレーターの解法の適用⇒ハートリー・フォック法

分子軌道 $\phi$ の微小変化に対する変分原理( $\delta E/\delta \phi=0$ )

ハートリー・フォックエネルギー  

$$E = 2\sum_{i=1}^{n} h_i + \sum_{i,j=1}^{n} (2J_{ij} - K_{ij})$$



V. A. Fock



#### 非線形方程式の解法=SCF法

演算子が固有関数である分子軌道 Øを含むため, ハートリー・フォック方程式は非線型方程式

<u>Ab initio ハートリー・フォックSCF法</u>

1 初期分子軌道 $\{\phi\}$  → フォック演算子*F* 2 フォック演算子*F* → ハートリー・フォック方程式 分子軌道 $\{\phi_i^{(1)}\}$  & 軌道エネルギー $\{\varepsilon^{(1)}\}$ 3  $\{\phi_i^{(1)}\}$  →  $F^{(1)}$  →  $\{\phi_i^{(2)}\}$  &  $\{\varepsilon_i^{(2)}\}$ 4  $\{\phi_i^{(n-1)}\} = \{\phi_i^{(n)}\}$  かつ $\{\varepsilon_i^{(n-1)}\} = \{\varepsilon_i^{(n)}\}$ なら終了 そうでないなら3へ戻る

## コンピュータによる計算のために

ハートリー・フォック方程式の実際的な解法:
 フォンノイマン型コンピュータの得意な行列計算に
 <u>ローターン法</u>
 分子軌道関数を基底関数の組{<sub>Xp</sub>}で展開

$$\phi_i = \sum_p^{N_{\text{basis}}} \chi_p C_{pi} = \chi \mathbf{C}_i$$



C. C. J. Roothaan



ハートリー・フォック方程式は行列形式に

$\hat{F}\phi_i = \varepsilon_i\phi_i$	$\Rightarrow$	$\mathbf{FC}_i = \varepsilon_i \mathbf{SC}_i$	F: フォック行列 S: 重なり行列
		$ \mathbf{F} - \varepsilon_i \mathbf{S}  = 0$	C <sub>i</sub> : 係数行列

実際には、規格直交な基底関数(S=E)を作成し、 フォック行列 F を対角化して軌道エネルギー ε<sub>i</sub>を求める

#### 電子相関の定義

<u>電子相関</u>

ハートリー・フォックエネルギーは全エネルギーの99.8%以上を与える
 ハートリー・フォック法では化学反応を定性的にも与えられない
 ⇒電子相関が足りないため必要な精度が得られないから
 化学反応を与えるのに必要な精度(化学的精度)
 =結合距離0.1Å程度,結合エネルギー数kcal/mol程度の誤差

<u>電子相関の定義</u> [Löwdin, Phys. Rev. <u>97</u>, 1509, 1955.] 電子相関エネルギーは厳密エネルギーと ハートリー・フォックエネルギーとの差

 $E_{\rm correlation} = E_{\rm exact} - E_{\rm HF}$ 



P.-O. Löwdin

#### 電子相関はなぜ生じる?

<u>電子相関が生じる理由</u> ハミルトニアン演算子とシュレーディンガー方程式は矛盾? ハミルトニアン演算子 シュレーディンガー方程式  $\hat{H} = \hat{h} + \sum_{i < i}^{N} \frac{1}{r_{ii}}$  $\hat{H}\Psi = E\Psi$ 

ハミルトニアン演算子は特異点(r<sub>ii</sub>=0)をもつのに, シュレーディンガー方程式の右辺は特異点をもたない



#### 動的電子相関

電子相関は、動的電子相関と静的電子相関の2種類にわけられる

ハミルトニアン演算子の特異点を取りのぞくための必要条件 波動関数の相関カスプ条件 -ロン孔(相関孔)

$$\frac{\partial \Psi}{\partial r_{12}}\Big|_{r_{12}=0} = \frac{\Psi}{2}\Big|_{r_{12}=0}$$

を満たさない(r<sub>12</sub>→0で左辺0)

⇒動的電子相関がたりない

ハートリー・フォック波動関数は相関カスプ条件 カスプ r<sub>12</sub> 図. クーロン孔

#### 動的電子相関

おもにハートリー・フォック法が余計に取りこんだ短距離電子間のクーロン反発

動的電子相関は、化学結合(反平行スピンの対)が切れるとつねに減る 動的電子相関が大きいならば, ハートリー・フォック法は 結合長を長く、結合エネルギーを小さく与える

# 静的電子相関

#### 静的電子相関

エネルギーの近い電子配置間の長距離相互作用(擬縮退効果)

静的電子相関は,結合が切れるときの軌道の入れかわりにおいて効く 静的電子相関が重要ならば,ハートリー・フォック法は 結合長を短く,結合エネルギーを大きく与える

結合が入れかわるとき、電子配置のエネルギー が近くなるために<mark>摂動展開の分母が小さくなり、</mark> 電子相関が大きくなる

静的電子相関を与えるには、エネルギーの近い 結合性軌道を組みかえた励起電子配置の線形 結合からなる多配置波動関数を作ればよい

きわめて高精度な方法として、多配置波動関数 をHF波動関数と置きかえて、動的電子相関を 取りこむ多参照法がある



ホーヘンベルク・コーン定理:基礎定理

密度汎関数法はホーヘンベルク・コーン定理にもとづく理論 [Hohenberg & Kohn, Phys. Rev. B, <u>136</u>, 864, 1964.]



ホーヘンベルク・コーンの定理



P. Hohenberg

第1定理:外場ポテンシャルは電子密度で決定可能
 W. Kohn
 電子密度は一意的に外場ポテンシャル
 (分子構造で決まる1電子ポテンシャル)を決定できることを証明



第2定理:電子密度であらわされた変分原理 (N表現可能な)電子密度であらわされた ハミルトニアン演算子は、必ず最低エネルギー をもつことを証明







# DFTが化学において有効である理由

化学=物質の電子状態間の電子のやりとりによる状態変化

化学を定量的に与えるには、電子相関をバランスよく取りこむことが必要

動的電子相関

相関カスプをもたないため余計に取りこまれた短距離電子間のクーロン反発

静的電子相関 エネルギーの近い電子配置間の長距離相互作用(擬縮退効果)

DFTの交換・相関汎関数はこの2種の電子相関をバランスよく取りこむ







局所密度近似(LDA):	<mark>電子密度</mark> ρのみであらわした汎関数
一般化勾配近似(GGA):	LDAを <mark>密度勾配▽</mark> ρで補正した汎関数
メタGGA(meta-GGA):	GGAを運動エネルギー密度τで補正した汎関数
混成GGA(hybrid-GGA):	GGAをハートリー・フォック交換積分を一定割合混合
	して補正した汎関数
半経験的汎関数:	多くの半経験的パラメータを使い,高精度に物性を
	与えることを目的とした汎関数
プログレッシブ汎関数:	いっしょに使う <mark>汎関数</mark> により形を変える汎関数

### 汎関数に対する1つの考えかた



図. ヤコビのはしご

普遍的汎関数を導くには、変数に 対する摂動項(のようなもの)を つみあげて近似を高めていけば いいという1つの考えかた [Perdew et al., J. Chem. Phys. <u>123</u>, 062201, 2006.]

# 量子化学におけるDFT汎関数開発



## 汎関数に対して最近白熱した議論

[Medvedev, Bushmarinov, Sun, Perdew, and Lyssenko, Science 355, 49, 2017; ibid. 356, 496c, 2017.]



固有値 (i.e., 軌道エネルギー) を比較すべき

 $\left[-\frac{1}{2}\nabla^2 + V_{\text{eff}}(\mathbf{r})\right]\varphi_i = \varepsilon_i \varphi_i$ 

80年間にわたり, 軌道エネルギーを再現できる 理論はなかった

## 交換汎関数の長距離補正

交換汎関数の長距離相互作用誤差の補正

[likura, Tsuneda, Yanai and Hirao, J. Chem. Phys. 115, 3540, 2001.]



短距離部分には一般的な交換汎関数をそのまま使用  

$$\left(a_{\sigma} = \mu/(2k_{\sigma}), k_{\sigma} = (9\pi/K_{\sigma})^{1/2} \rho_{\sigma}^{1/3}\right)$$
  
 $E_{x}^{sr} \equiv -\frac{1}{2} \sum_{\sigma} \int d^{3}\mathbf{r} \rho_{\sigma}^{4/3} K_{\sigma} \left\{1 - \frac{8}{3}a_{\sigma} \times \left[\sqrt{\pi} \operatorname{erf}\left(\frac{1}{2a_{\sigma}}\right) + \left(2a_{\sigma} - 4a_{\sigma}^{3}\right) \exp\left(-\frac{1}{4a_{\sigma}^{2}}\right) - 3a_{\sigma} + 4a_{\sigma}^{3}\right]\right\}$ 

長距離部分はハートリー・フォック交換積分で補正

$$E_{\rm x}^{\rm lr} = -\frac{1}{2} \sum_{\sigma} \sum_{i}^{\rm occ} \sum_{j}^{\rm occ} \int \int d^3 \mathbf{r}_1 d^3 \mathbf{r}_2 \phi_{i\sigma}^{*}(\mathbf{r}_1) \phi_{j\sigma}^{*}(\mathbf{r}_2) \frac{\operatorname{erf}(\mu r_{12})}{r_{12}} \phi_{j\sigma}(\mathbf{r}_1) \phi_{i\sigma}(\mathbf{r}_2)$$

長距離補正をもとに、さまざまな汎関数が開発されてきた:

CAM-B3LYP [Yanai, Tew, & Handy, Chem. Phys. Lett. <u>91</u>, 551, 2004.] LC-@PBE [Vydrov, Heyd, Krukau, & Scuseria, J. Chem. Phys. <u>125</u>, 074206, 2006.] BNL [Livshits & Baer, Phys. Chem. Chem. Phys. <u>9</u>, 2937, 2007.] @B97-series [Chai & Head-Gordon, J. Chem. Phys. <u>128</u>, 084106, 2008.] Tuned RSH [Baer, Livshits, & Salzner, Ann. Rev. Phys. Chem. <u>61</u>, 85, 2010.]

## 長距離補正が解決・改善した問題





#### 軌道エネルギーの物理的意味

#### 分数占有数に対するエネルギー直線性定理 [Perdew et al., Phys. Rev. Lett. 49, 1691, 1982; Yang et al., Phys. Rev. Lett. 84, 5172, 2000.] 全電子エネルギーは占有数の変化 $E\left(n+\frac{p}{a}\right) = \frac{p}{a}E(n+1) + \frac{q-p}{q}E(n)$ について直線的に変化する ヤナクの定理 [Janak, Phys. Rev. B 103, 7165, 1978.] E(n) 1 ある分子軌道の占有数に関する $\partial E$ $= \mathcal{E}_i$ 全電子エネルギーの変化は $\partial n_i$ 軌道エネルギーに一致する IP $\frac{\partial E}{\partial n}(\Delta n = +0)$ エネルギー直線性定理+ヤナクの定理 = 一般化クープマン定理 $= \mathcal{E}_{LUMO}$ $\frac{\partial E}{\partial n} \left( \Delta n = -0 \right)$ EA HOMOエネルギー ε<sub>HOMO</sub> = -IP ∂n **LUMOエネルギー** ε<sub>LUMO</sub> = -EA = *Е*<sub>НОМО</sub> シャム・シュリューター定理 [Sham and Schlüter, Phys. Rev. B 32, 3883, 1985.] -1 0 最外殻軌道エネルギーは占有数と無関係 +1 n

分子軌道のイメージング

図. CO<sub>2</sub>分子のHOMOの一般化 分子軌道トモグラフィ像 [Vozzi et al., Nature Phys. <u>7</u>, 822, 2011]



分子軌道イメージング

 [Itatani et al., Nature <u>432</u>, 867, 2004]:
 分子のトモグラフィスキャン

 レーザーをすべての方向から制御し各軌道 から電子を脱離させて衝突角を測定

#### 図. PTCDA分子の分子軌道 [Puschnig et al., Phys. Rev. B 7, 235427, 2011]



#### 応答物性

n <sub>F</sub>	n <sub>B</sub>	n	n <sub>R</sub>	応答物性	一化	学で	対象	となる	物性のほとんどは応答物性	
0	0	0	0	エネルギー	/Ľ	合物	リモロ	エጥ/	レキーの偏似分	
1	0	0	0	双極子モーメント			広公	地加州士	$\sim \partial^{n_F+n_B+n_I+n_R} E$	
0	1	0	0	磁気双極子モーメント	心實物注 $ \sim \frac{\partial \mathbf{F}^{n_F} \partial \mathbf{B}^{n_B} \partial \mathbf{I}^{n_I} \partial \mathbf{R}^{n_R}}{\partial \mathbf{F}^{n_F} \partial \mathbf{B}^{n_B} \partial \mathbf{I}^{n_I} \partial \mathbf{R}^{n_R}}$					
0	0	1	0	超微細結合定 <b>数</b>						
0	0	0	1	エネルギー勾配	<sup>II</sup> F	П <sub>В</sub>	"	"R	心合物に	
2	0	0	0	<b>公</b> 極來	1	0	0	1	赤外 (IR) 吸収強度	
2	2	0	0	ガルを	1	1	0	0	旋光性,円偏光二色性	
U	Z	0	0	1XX1C <del>~</del>	0	1	1	0	核磁気遮蔽	
0	0	2	0	核スピン・スピン結合	2	0	0	1	ラマン強度	
0	0	0	2	調和振動子	— 3	0	0	1	超ラマン効果	
					2	1	0	0	磁気円偏光二色性 (ファラデイ効果)	
表. 応答物性と導関数の次数 [Jensen, "Introduction to Computational Chemistry, 2 <sup>nd</sup> Ed. (Wiley, Chichester, UK), 2007.]より抜粋			1	0	0	2	倍音や結合バンドに対する 赤外強度			
			2	0	0	2	倍音や結合バンドに対する ラマン強度			
					2	2	0	0	コットン・ムートン効果	

### 応答物性計算法

<u>例:外部電場の応答物性</u>

## Diels-Alder反応におけるハードネスの変化



固有反応座標

軌道エネルギーにもとづく反応解析

[Tsuneda & Singh, J. Comput. Chem. <u>35</u>, 1093, 2014.] 定量的軌道エネルギーにもとづく反応解析法





# マリケンポピュレーション解析は正しいか?

図. 実際の電子移動方向とマリケンポピュレーション解析の結果の比較

Hydrogen	Transfer Rea	actions	Heavy-Atom Tran	sfer Reactions	Sigmatro	pic Shifts	
(1) H + HCl	<>	HCl + H	(21) N O + H	N + OH	(36)	(	
(2) OH + H <sub>2</sub>	<>	$H_2O + H$	(21) N <sub>2</sub> O + II	$\sim N_2 + OII$			
(3) CH <sub>4</sub> + H	<b>~~~</b> >	$\mathrm{CH}_3 + \mathrm{H}_2$	(22) $H + FH$	→ HF + H	(37)	$\leftrightarrow$	順反応で
(4) $CH_4 + OH$	←→	$\mathrm{CH}_3\ +\mathrm{H}_2\mathrm{O}$	(23 $H + FCH_3 \checkmark$	→ HF + CH	3		详
(5) CH <sub>3</sub> OH +	н 🔶 🕨	$\rm CH_2OH + H_2$	Nucleophilic Subst	itution Reactions	(38)	→	
(6) H <sub>2</sub> + H	<>	$H + H_2$	$(24) CH_3F + F^- \checkmark$	$\rightarrow$ FCH <sub>3</sub> + F			
(7) $NH_2 + OH$	<b></b>	$NH_2 + H_2O$	(25) $F + CH_3Cl \blacktriangleleft$	$\rightarrow$ CI + CH <sub>3</sub> F	Cycl	loaddition	_
		$CH \pm HCl$	(26) $OHCH_3 + F^- \blacktriangleleft$	$\rightarrow$ OH <sup>-</sup> + CH <sub>3</sub>	F (39)		
(8) $CH_4 + CI$	~~~~>	CH <sub>3</sub> + HCI	(27) Cl <sup>-</sup> CH <sub>3</sub> Cl ◀	► CICH <sub>3</sub> C	-		逆反応で
(9) $C_2H_6 + OH$	<b>←</b>	$C_2H_5 + H_2O$	Unimolecular and A	ssociation Reaction	ons (40)		2米
(10) $H_2 + F$	<>	HF + H	(28 HN <sub>2</sub> <	$\rightarrow$ H + N <sub>2</sub>			迓
(11) $CH_3 + OH$	<>	$\mathrm{CH}_4 + \mathrm{O}$	(29) HCO <	→ CO + H	(41) 2	$\leftrightarrow$	$\searrow$
(12) PH <sub>3</sub> + H	<b>~ &gt;</b>	$PH_2 + H_2$	(30) $H + C_2 H_4$	→ CH <sub>3</sub> CH <sub>2</sub>		$\sim$	~/
(13) H + OH	<b></b>	$H_2 + O$			Cyclorev	ersions	
(14) $H_2S + H$	←→	$\mathrm{SH} + \mathrm{H}_2$	(31) $\operatorname{Ch}_3 + \operatorname{C}_2 \operatorname{H}_4$		(42)	( <sup>\</sup>	両方向で
(15) CI + OH	<b></b>	HCl + O	Electrocyclic Read	ctions	$\triangle$		/ 逆
				1			
(16) $NH_2 + CH_3$	~~~>	$NH + CH_4$	(33)			←→ (	
(17) $NH_2 + C_2H_2$	5 <b></b>	$NH + C_2H_6$		$\wedge$		$\rangle$ $\backslash$	/
(18) $NH_3 + C_2H$	5 🔶 🗡	$\mathrm{NH}_2 + \mathrm{C}_2\mathrm{H}_6$	(34)	240	% の2 体 反 [	ふについて	峃
(19) $NH_2 + CH_4$	<b>~~~</b> >	$\rm NH_3 + CH_3$	~ //	24.			
(20) cis-C <sub>5</sub> H <sub>8</sub>	←→	$cis-C_5H_8$		O→	反心性軌〕 マリケンポ	直以外の朝 パッレーショ	」 迫の奇与 コンの限界
					× ) / / ///		

## 有機電子論の理論的裏付け

反応軌道エネルギーダイアグラムは経験的な有機電子論を物理的に基礎づける



非占有反応性軌道は反応の進行による占有反応性軌道の変化で偏った 電子分布を回復するのに利用

有機電子論は反応過程での占有反応性軌道を表現
▶ ダイアグラムは有機電子論の理論的背景となりえる

#### データ科学としてのDFT

[Tsuneda, The Chemical Record, in press.]

<u>データ科学における予測モデルの決定プロセス</u>

- 1. 予測に寄与する記述子を収集
- 2. パラメータを使ったデータ科学的モデルで予測関数を決定
- 3. 汎化性能を上げるために
   特徴量の種類を増やしてモデルを洗練化
   →過学習を避けるために
   特徴量を削減

#### DFT汎関数の開発プロセス

- 1. 系の電子状態に寄与する物理的効果を収集
- 2. 半経験的パラメータを使った物理的効果を結合したモデルで汎関数を決定
- 適用可能性を改善するために物理的効果の種類を増やして汎関数を洗練 化→ポテンシャルの偽の極小点のような問題を避けるため、半経験的パラ メータの数を削減

DFT汎関数開発は化学において最も成功したデータ科学

# データ科学のDFTへの応用

[Rupp, von Lilienfeld & Burke, J. Phys. Chem. <u>148</u>, 241743, 2018.]



## 機械学習DFT開発の要件

[Tsuneda, The Chemical Record, in press.]

#### 機械学習DFT開発にあたっての質問事項

- 1. 記述子に何を選べばいいか?
- 2. 導入されるパラメータには何が要求されるのか?
- 3. 汎関数を物理的に規定する基礎物理条件の予測精度への影響は?
- 4. 予測関数としてのDFT汎関数は何をターゲットにすべきか?

#### <u>質問への回答</u>

- できるだけ独立な記述子→汎化性能を向上させるため 物理的意味をもつ記述子→結果の解析を用意にするため
- 2. 最少のパラメータ→ポテンシャル上の偽の極小点の問題を回避するため 物理的意味をもつパラメータ→回帰係数で結果を解釈できるように
- 3. 特に運動エネルギーで基礎物理条件の影響は大きい 基礎物理条件を使うとパラメータ削減できる
- 4. 物性値のベンチマークセットの平均絶対誤差の最小化は× 原子化エネルギーは分子サイズ依存するので×

# より詳しく知りたい方へ



<u>"常田,『密度汎関数法の</u> 基礎』(講談社, 2012)"



<u>英語版も</u> <u>"T. Tsuneda, Density Functional</u> <u>Theory in Quantum Chemistry</u> <u>(Springer, 2014)"</u>