

水素の先端計算による水素機能の高精度解析

常 行 真 司*

1. はじめに

ユビキタスな元素である水素は、様々な物質中に不純物として存在し、半導体の不動態化(注: キャリアドープした半導体の電氣的活性が失われること)を引き起こす例のように、原子周りの環境に応じて異なる荷電状態をとることで、物性を大きく左右することが知られている。また近年、様々な酸化物中の酸素をヒドリドで大量に置換してキャリアドープできることや、新しいヒドリド伝導体、高温超伝導を示す超高压下の硫化水素などが見つかり、固体中の水素が改めて脚光を浴びている。一方、電極反応や各種触媒反応、合金の水素吸蔵・放出など、水素社会実現のために必要な化学反応も、近年さらに重要性が増している。こういった水素を含む物質・材料の研究では、水素量や水素原子位置、ダイナミクスの計測の難しさが、常に大きな問題となっている。

新学術領域研究「ハイドロジェノミクス: 高次水素機能による革新的材料・デバイス・反応プロセスの創成」の計画研究 A05-2「水素の先端計算による水素機能の高精度予測」では、この水素の問題を計算科学の立場で解決するため、第一原理シミュレーション手法の開発を行っている。中でも、計測とシミュレーションを直接つなぐデータ同化手法の開発は、本領域に特徴的な試みである。

本稿では、まず第2章で物質・材料の研究を支える第一原理シミュレーション手法の現状を紹介する。とくに水素で顕著な原子核の量子効果、有機物のシミュレーションで重要なファンデルワールス相互作用、水素化合物の超伝導転移温度の定量計算は、ハイドロジェノミクス研究を進める上で重要な手法である。第3章では、実験的観測が困難な水素を含む物質・材料のデータ同化構造探索手法について述べる。計算機シミュレーションだけを使って物質の結晶構造を予測

することは、原子数が増えるにつれて可能性の数が爆発的に増えるため、非常に困難になる。そこで可能性を絞るために実験データを利用しようというのが、我々の提案するデータ同化手法である。

2. 水素を含む系の第一原理シミュレーション

(1) 電子状態計算

物質・材料の結晶構造、電子状態、物性のシミュレーション手法のうち、周期系、液体、表面・界面といった凝縮系一般で経験パラメータを使わないシミュレーションを可能にするのが、密度汎関数理論(Density Functional Theory, 以下DFT)に基づく第一原理電子状態計算と、そこから計算される原子間力を使って原子を動かす第一原理分子動力学法(First-Principles Molecular Dynamics, 以下FPMD)である。

DFTは1960年台にW. Kohnらが作り上げた理論で、物質の基底状態エネルギーが電子密度分布だけで決まり、その計算に必要なエネルギー汎関数は物質によらないユニバーサルなものであるということを示したHohenberg-Kohnの定理⁽¹⁾と、電子密度の計算を相互作用しない仮想的な電子系の問題にマップすることで、電子密度とエネルギーの精密な計算手法を与えたKohn-Sham理論⁽²⁾からなる。DFTにより、波動関数理論ではできなかったような凝縮系の電子状態計算が可能となり、開発者のKohnは1998年にノーベル化学賞を受賞した⁽³⁾。

DFTの実際の計算では、電子密度からエネルギーを計算する際に、何らかの近似に基づく交換相関エネルギー汎関数を利用する。有名なのは一様電子ガスの結果を利用する局所密度近似(LDA)や密度勾配の補正を加えた一般化密度勾配近似(GGA)であるが(文献(3)およびその引用文献を参照)、最近では、波動関数理論であるハートリー・フォック法の交

* 東京大学大学院理学系研究科; 教授(〒302-0124 東京都文京区本郷7-3-1)
Hydrogen Function Analysis by Advanced Computer Simulations; Shinji Tsuneyuki (School of Science, The University of Toyo, Tokyo)
Keywords: *hydrogenomics, data assimilation, first-principles simulation, density functional theory (DFT), molecular dynamics (MD)*
2021年1月25日受理[doi:10.2320/materia.60.176]

換エネルギーの表式を取り入れたハイブリッド汎関数もよく用いられている⁽⁴⁾。また局在性の強いd軌道やf軌道に、電子相関効果を格子モデル的に取り入れるDFT+U法も使われている⁽⁵⁾。

これらの汎関数では、極性のない分子間の相互作用として重要なファンデルワールス(vdW)相互作用の記述が不十分で、分子性の固体、液体、表面分子吸着系などの計算結果の精度が悪い。そこで近年、簡便で正確にvdW相互作用を計算するための手法が研究されている。濱田により提案されたvdW密度汎関数は、その中でも成功を収め、広く使われ始めた手法である⁽⁶⁾。

(2) 動力学計算

1985年になると、原子の動きとそれに伴う電子状態変化を同時に解くことで計算量や計算に必要な計算機のメモリー量を大幅に減らすことのできるCar-Parrinello法が登場し、液体や表面の動力学シミュレーションもDFTで行えるようになった⁽⁷⁾。この手法を電子状態の逐次解法へと発展させ、より精密な電子状態計算と分子動力学計算を実現したのが、現在のFPMDである⁽⁸⁾。FPMDは動力学計算だけでなく、ポテンシャルエネルギー面をたどる構造最適化にも利用される。

近年、電極での水素発生のように、電子のリザーバーである固体表面と外部のイオンの間で電子移動が起きる化学反応を、FPMDで調べたいというニーズが高まってきた。そのような電子移動を伴うシミュレーションを有限サイズの系で行うと、電子1個の移動で表面が無視できないほど帯電し、電位が変化するという問題が生じる。大谷と杉野による有効遮蔽媒質法は、この問題を解決して電子数一定ではなく電位一定条件という開放系のシミュレーションを実現する手法であり、FPMDと組み合わせて広く利用されるようになってきている⁽⁹⁾。

(3) 原子核の量子効果

FPMDでは、与えられた原子位置を使ってDFTに基づく電子状態計算を行い、そこから原子に働く力を計算し、ニュートンの運動方程式を使って原子を動かす。すなわち原子は古典的な粒子として扱っているの、しばしば用いられる「量子MD」という言い方は誤解を生みやすく推奨できない。これに対して、本稿の主題である水素原子は質量が軽いいため、固体中にある時の典型的な零点振動エネルギーが0.1 eV程度となり、これは温度に換算すると1000 K以上の大きな値である。つまり室温程度の温度では、分布関数の広がりやの主な要因は熱揺らぎではなく量子効果ということになる。このような、エネルギーや分布に現れる量子効果を、我々は通常無視することが多いが、本当に無視して良いかどうかは物質や現象ごとによく検討しなければならない。

量子効果が重要と思われる場合には、ファインマンの経路積分法を使うことで、量子力学的な原子をバネでつながれた複数の仮想古典粒子として近似的に記述し、分子動力学法

(MD)を使って温度効果も含めてシミュレーションすることができる。この手法を経路積分分子動力学法(Path-Integral Molecular Dynamics, 以下PIMD)と言い、とくに第一原理に基づくものをFP-PIMDと言う。FP-PIMDはD. Marxらによって初めて水素を含む分子に適用され⁽¹⁰⁾、その後、著者のグループが初めて凝縮系(結晶シリコン中の不純物水素⁽¹¹⁾、超高圧下の固体水素⁽¹²⁾)に適用して、原子核の量子効果を論じた。

PIMDでは、量子系を古典系にマップするにあたって、計算量が仮想粒子の数の分だけ増える。量子効果が大きい場合ほど、近似の精度を担保するのに必要な仮想粒子の数を増やす必要があり、水素を含む系の典型的な計算量は通常のMDの100倍のオーダーとなる。そのため、もともと計算量の多いDFTを使ったFP-PIMDの応用はなかなか広がらなかった。幸いFP-PIMDは並列計算しやすいため、近年の並列計算機の発展に伴って利用しやすくなり、日本では志賀らが非常に汎用性の高いプログラムを開発・公開したことで、さまざまな系に应用が広がっている⁽¹³⁾。

(4) 超伝導

第一原理に基づく超伝導転移温度の計算には様々な報告例があるが、多くの場合、DFTによる第一原理電子状態計算と電子格子相互作用を表す経験パラメータを組み合わせて用いており、予言力には限界があった。近年、ドイツのE.K.U. Grossらが、DFTを拡張して超伝導転移を記述する理論(超伝導密度汎関数理論、以下,SCDFT)を提案し、少なくともフォノンを媒介とする超伝導については、真に非経験的な転移温度予測の道が開かれた⁽¹⁴⁾。その後フォノン以外、たとえばプラズモン⁽¹⁵⁾やスピン揺らぎ⁽¹⁶⁾の効果を取り入れる試みもなされている。

固体水素は以前から圧力により金属化すると予想され、金属化すれば高温超伝導体になるであろうと考えられてきた。残念ながら固体水素の金属化が起きる圧力はあまりにも高く、水素単体での超伝導は実現されていないが、水素を多量に含む系で高温超伝導の発現がいくつか実証されている。そのきっかけとなったのは、ドイツのマックスプランク研究所のグループが発見した、硫化水素高圧相での200 Kを超える温度での超伝導転移である⁽¹⁷⁾。このような水素系の超伝導転移は、基本的にはフォノン媒介であると考えられ、実際SCDFTを使うと転移温度がよく再現される⁽¹⁸⁾⁽¹⁹⁾。その時、水素が軽く量子効果による分布の広がりが大きいことを反映し、格子振動の非調和効果が大きいことも指摘されている⁽²⁰⁾。

3. データ同化構造探索

これまで紹介した第一原理シミュレーション手法では、いずれの場合も初期原子座標を入力として与え、そこからエネルギーが下がるように原子を動かしたり、温度に見合った初速度を与えて動力学計算をしたりする。では適切な初期原子

座標がわからない場合、どうすれば良いだろうか。

典型的な例は、化学組成から未知の結晶構造をシミュレーションで探索したいという場合である。そのような場合には、たとえばランダムな初期原子配置をたくさん用意してそれぞれ構造最適化を行い、最も安定なものを選ぶ(Random Sampling)、液体のように原子が動き回れる高い温度のMDから始めて徐々に温度を下げて安定構造を見つける(Simulated Annealing, 以下SA)、準安定構造の一つ見つけたら、その近傍でポテンシャルエネルギー障壁を越えるように原子を動かして次の安定点を見つける(Basin Hopping, Minima Hopping)、といった様々な手法が提案されている。また最近成功を収めている手法には、ランダムな初期配置の構造最適化に、遺伝子のような交叉と突然変異を組み合わせた遺伝アルゴリズム(Genetic Algorithm)や、鳥が群れになって餌場を探すかのように多次元空間を探索する粒子群最適化法(Particle Swarm Optimization)がある[レビュー論文として、文献(21)をあげておく]。これらの結晶構造探索は、シミュレーションセルに含まれる原子の座標からなる多次元空間での最適化問題であり、正しい解を得ることは原子数(次元)が増えるにつれて急激に困難になる。現状では100原子くらいになると、もはや絶望的であろう。

そこで我々が注目したのは、データ同化の考え方である。気象データは離散的な点でしか得られないため、気象モデル(運動方程式)を解くための初期値を内挿で決めるしかない。また気象モデルも完全なものではない。そのため、時間が経てばシミュレーション結果と現実のずれが大きくなるので、時間と共に次々と入ってくる観測データを使って、シミュレーションを改善しようと考えた。これが気象予報の分野で長年にわたり様々な手法が開発されてきた、データ同化の基本的なアイデアである。

結晶構造予測の場合には、時系列の実験データではないが、手に入れやすいX線や中性子線の粉末回折実験データを使う。もちろん良質のサンプルがあってきれいな回折実験データが手に入るならば、第一原理に基づく計算機シミュレーションは必要ない。データ同化が役に立つのは、試料の質が悪かったり量が少なかったり、あるいはダイヤモンドアンビルセルを用いた超高压実験のように実験装置の制限で回折角に限られるなどの理由から、回折実験データだけでは構造が決まらない場合である。水素のようにX線ではほとんど見えない元素の位置を決めたい場合にも有用である。以下では回折実験データを用いたデータ同化構造探索手法について、少し詳しく説明しよう。

全原子の位置をまとめて R 、粉末回折実験データを $I(\theta)$ とシンボリックに書くことにする。 I は回折角 θ の関数である。有限温度での原子の分布関数を $\rho(R)$ 、 I の分布関数を $\rho(I)$ 、実験から I が与えられたときの原子の分布関数を $\rho(R|I)$ 、逆に R がわかったとき I が観測される確率分布を $\rho(I|R)$ とすると、これらの間には

$$\rho(R|I) = \frac{\rho(I|R)\rho(R)}{\rho(I)} \quad (1)$$

というベイズの定理が成り立つ。有限温度 T でのカノニカル分布を考えると、右辺の $\rho(R)$ は、ポテンシャルエネルギー $E(R)$ と温度の逆数 $\beta(=1/k_B T)$ を用いて、

$$\rho(R) \propto \exp[-\beta E(R)] \quad (2)$$

と書ける。 R が決まれば回折ピークの角度や大きさはすべて計算できるので、その計算値を $I_{\text{calc}}(\theta; R)$ とする。 $\rho(I|R)$ を考える際には、 θ の関数である $I_{\text{calc}}(\theta; R)$ が $I(\theta)$ からずれるにつれてその確からしさが減ると考え、ずれが大きいほど大きなペナルティーを与えるようなエネルギーの次元をもった汎関数 $D[I, I_{\text{calc}}]$ を導入して、

$$\rho(I|R) \propto \exp(-\beta D[I, I_{\text{calc}}]) \quad (3)$$

と書く。これらをベイズの定理の式に代入すると、

$$\rho(R|I) \propto \exp\{-\beta(E(R) + D[I, I_{\text{calc}}])\} \quad (4)$$

が得られる。構造シミュレーションの各ステップでの R がわかればその構造に対する I_{calc} が計算され、そこからペナルティー D も決まるので、 $D[I, I_{\text{calc}}]$ を明示的に $D(R)$ と書くことにしよう。結果的に我々が欲しい $\rho(R|I)$ は、仮想的なポテンシャルエネルギー関数

$$F(R) \equiv E(R) + D(R) \quad (5)$$

のもとでの原子分布関数であることがわかる。構造探索のもっとも簡単な手法の一つであるSAでは、 $E(R)$ で相互作用しながら運動する原子のシミュレーションを融点より高い温度で開始し、絶対零度まで徐々に温度を下げていくことで最安定構造を探す。回折ピークに関して何らかの情報がある場合には、 $E(R)$ に替えて $F(R)$ を使ったSAを行えば、その情報を利用できるわけである。

このことをもう少し直感的に示したのが、図1である。第一原理シミュレーションで得られる $E(R)$ は、一般に非常に複雑な多谷構造をもち、SAなどの手法で最安定点を見つけようとしても準安定点にトラップされて構造探索に失敗する。ペナルティー $D(R)$ はこれとは異なる多谷構造を持つと考えられるが、最安定構造(図1の星印)は $E(R)$ と $D(R)$ に

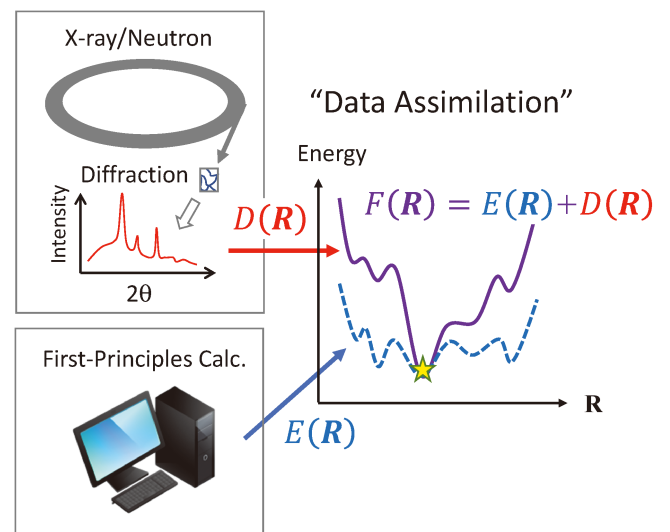


図1 データ同化構造探索手法の模式図。(オンラインカラー)

共通の安定点なので、両者をたしあわせた $F(R)$ ではその安定性が強調され、探索が容易になる⁽²²⁾。

ペナルティ関数は人為的なものなので、様々な選び方がある。素直に考えれば、観測された回折データ $I(\theta)$ と計算された回折データ $I_{\text{calc}}(\theta; R)$ の相関係数(を1から減じたもの)を利用すると良いが、たとえば回折ピークの現れる回折角だけで定義される結晶化度 λ を使うこともできる⁽²²⁾。

$$D(R) = D[I(\theta), I_{\text{calc}}(\theta; R)] = \alpha N(1 - \lambda) \quad (6)$$

$$\lambda = \frac{\sum_{\theta_{\text{obs}} - \delta}^{\theta_{\text{obs}} + \delta} I_{\text{calc}}(\theta; R) d\theta}{\int_{\theta_{\text{min}}}^{\theta_{\text{max}}} I_{\text{calc}}(\theta; R) d\theta} \quad (7)$$

式(6)の α はエネルギーの次元を持つ定数で、データ同化の重みを決めるコントロールパラメータ、 N はシミュレーションに用いる原子数である。式(7)で定義される λ の分母は、実験で計測された角度範囲 $\theta_{\text{min}} \leq \theta \leq \theta_{\text{max}}$ で $I_{\text{calc}}(\theta; R)$ を角度積分した規格化因子、分子は同じく $I_{\text{calc}}(\theta; R)$ の積分であるが、積分範囲は実験で回折ピークが観測される角度 θ_{obs} の近傍(幅 2δ) のみに限定する。こうすれば常に $0 \leq \lambda \leq 1$ であり、 $\theta_{\text{min}} \leq \theta \leq \theta_{\text{max}}$ に観測されたすべての回折ピークの回折角が(強度を別にして)計算値と一致すれば、 $\lambda = 1$ ($D = 0$) となる。ポテンシャルエネルギーに対するペナルティの重みを決めるパラメータ α は、何らかの原理により決めることができないが、経験的には融点の熱エネルギーの10倍程度でうまくいくようである。

実際にこのデータ同化手法を用いてシリコンの構造探索を行った時の構造およびそこから計算される X 線回折データの変化の様子と、その過程で計算された $E(R)$ (破線) と $F(R)$ (実線) を、**図2** に示す。この計算は原子数64個の周期セルと原子間力モデル⁽²³⁾ を用い、古典 MD のプログラムパ

ッケージ LAMMPS⁽²⁴⁾ を使って初期温度5000 K から 0 K まで 5 ps かけて SA を行った。左下図の横軸は結晶化度 λ である。SA の時間発展は概ね左から右に進み、この例では最終的に $\lambda = 1$ となって完全結晶にたどりついている。データ同化に使ったのは、回折角60度以下に見られる3本のピーク位置のみである。初期構造はランダムなので、回折ピークは周期的境界条件が許すすべての角度に現れる。時間が経つにつれて、データ同化に使った3箇所の正しい回折ピーク(×印のついた線)と同じ角度のピークが成長し、正しい結晶構造が得られる。この過程で、 $E(R)$ の変化を見ると活性障壁を超えているように見えるが、計算に使った $F(R)$ の方は単調に減少しており、これにより $E(R)$ を使う SA よりも短時間で結晶構造に到達したことがわかる。

シリコンは結晶構造が簡単だが、それでも通常の SA でこれほど急冷してしまうと、結晶構造を得るのは容易ではない。対してデータ同化で正しい結晶構造が得られる確率(成功率)は、ほぼ100%であった。もちろん、このようなデータ同化がうまくいくかどうかは、コントロールパラメータ α の大きさ、SA の初期温度やどのくらいゆっくりと温度を下げるかによって変わる。化学組成や凝集機構、構造の複雑さ、実験データの質と量にも依存してはいる。したがって、未知物質の構造探索を行う場合、条件設定のための試行錯誤は避けられない。しかしながら我々がこれまで様々なタイプの結晶(酸化物、水酸化物、水素化物、合金、金属間化合物、イオン伝導体など)でテストした限りでは、いずれの場合もデータ同化によって構造探索の成功率が大きく向上することが確認された。たとえば高温超伝導を示す硫化水素の高圧相の場合、X 線回折実験で明瞭に観測されたピークは4本し

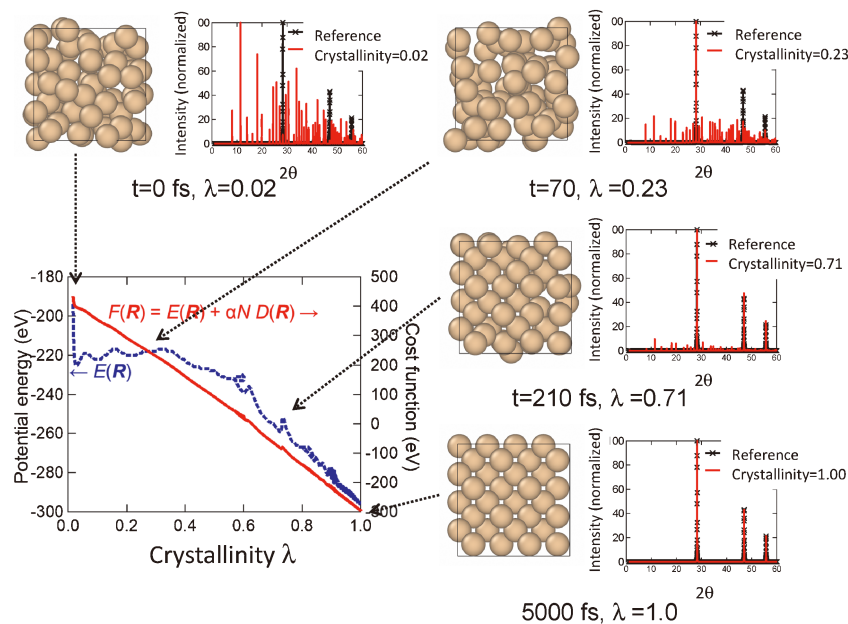


図2 シリコンのデータ同化構造探索シミュレーションで得られた原子配列の時間変化と、そこから計算される角度範囲 $0 < 2\theta < 60^\circ$ の粉末 X 線回折データ(赤)。データ同化に用いた既知の回折データを、×印つきの線(黒)で表示した。結晶化度 λ は時間と共に増加し、完全結晶の値である1に近づく。左下図では、その過程で得られた $E(R)$ (破線) と $F(R)$ (実線) を、結晶化度 λ に対してプロットした。(オンラインカラー)

