

科学研究費助成事業 研究成果報告書

令和 4 年 5 月 1 日現在

機関番号：14301

研究種目：若手研究

研究期間：2019～2021

課題番号：19K14680

研究課題名（和文）機械学習を用いた原子データインフラストラクチャの開発

研究課題名（英文）Development of Atomic Data Infrastructure with Machine Learning

研究代表者

藤井 恵介 (Fujii, Keisuke)

京都大学・工学研究科・助教

研究者番号：10637705

交付決定額（研究期間全体）：（直接経費） 3,100,000円

研究成果の概要（和文）：A) 多電子原子のエネルギー準位について、機械学習技術を用いた補完・品質評価を行った。特に、同電子系列と呼ばれる、電子数が同じで核電荷が異なる原子イオンのエネルギー準位が核電荷に対してスムーズな関数になることを利用し、その系列がどの程度連続的に繋がっているかを調べた。B) 未同定の発光線から、新たなエネルギー準位を発見する方法論の開発を行った。特に、千以上の発光線を用い、統計的な手法で効率よく準位同定を行うことを目指した。C) 大量エネルギー準位・遷移が存在し、それらの正確な数値計算が難しい多電子原子について、少量の原子データのみからその挙動を推定するモデルを構築した。

研究成果の学術的意義や社会的意義

宇宙プラズマや核融合プラズマの診断のため、鉄イオンやシリコンイオンの原子データのセットを用いた分光診断が必要であった。これまで、個別に報告される実験・理論データの収集、丁寧な品質評価、未知のデータの補完を行うことで、各分野において高品質な原子データセットが構築されてきた。しかし研究が進むにつれ常に新たな対象に関する原子データが必要になる。新たな対象に対しては十分に高品質なデータセットが構築されるまで、高精度な分光診断が実施できていない状況である。本研究で開発した機械学習法を応用することにより、個別に報告される原子データの妥当性の評価、新たな原子データ導出が容易になった。

研究成果の概要（英文）：A) We have carried out the data validation for the energy levels of many-electron atoms. In particular, so-called isoelectronic sequence, where the electron configurations are the same but only the nuclear charge is different, are analyzed. Based on the fact that the energy levels in an isoelectronic sequence follow a smooth function, erroneous data points are automatically found. B) We have developed a method that automatically finds a new energy levels from thousands unidentified emission lines observed from a many-electron atom. C) We have constructed a theoretical model that predicts the spectral properties of many-electron atoms.

研究分野：原子データ

キーワード：原子構造 統計物理

1. 研究開始当初の背景

宇宙プラズマや核融合プラズマの診断のため、鉄イオンやシリコンイオンの原子データ（各準位のエネルギー値やその間の遷移レート係数）のセットを用いた分光診断が必要であった。

個別に報告される実験データ・理論データの収集、丁寧な品質評価、未知のデータの補完を行うことで、各分野において高品質な原子データセットが構築されてきた。

近年、核融合炉壁にタングステンが使われることに変わった。また、宇宙プラズマではイットリウム、極端紫外光源プラズマではスズやハフニウムなどの分光診断が新たに必要になっている。

これらの新規データセットの構築も取り組まれているが、十分に高品質なデータセットが構築されるまで、高精度な分光診断が実施できていない状況である。このように、新たな元素・価数のイオンに関する診断が必要になる度、多大な時間・労力が費やされている。

2. 研究の目的

分光診断に際限なく必要な原子データセットを、機械学習技術により抜本的に効率よく構築することが本研究の目的である。特に当初は、

(2-1) 原子番号・価数・電子配置という離散変数を連続空間に埋め込む技術の開発

(2-2) 個別に報告されるデータの品質評価、欠損データの補完

(2-3) 全体の品質向上に本質的なデータの決定とその新規取得

を行うことを目指した研究計画であった。

3. 研究の方法

本研究課題は、主に以下の項目に関して研究を行った。

A) 多電子原子のエネルギー準位について、機械学習技術を用いた補完・品質評価を行った。特に、同電子系列と呼ばれる、電子数が同じで核電荷が異なる原子イオンのエネルギー準位が核電荷に対してスムーズな関数になることを利用し、その系列がどの程度連続的に繋がっているかを調べた。

B) 未同定の発光線から、新たなエネルギー準位を発見する方法論の開発を行った。特に、千以上の発光線を用い、統計的な手法で効率よく準位同定を行うことを目指した。

C) 大量エネルギー準位・遷移が存在し、それらの正確な数値計算が難しい多電子原子について少量の原子データのみからその挙動を推定するモデルを構築した。

4. 研究成果

A) 多電子原子のエネルギー準位の補完・品質評価

National Institute for Standards and Technology の Atomic Spectra Database (NIST-ASD) に収録されている多電子原子のエネルギー準位について、機械学習技術を用いた補完・品質評価を行った。特に、同電子系列と呼ばれる、電子数が同じで核電荷が異なる原子イオンのエ

エネルギー準位が核電荷に対してスムーズな関数になることを利用し、その系列がどの程度連続的に繋がっているかを調べた。

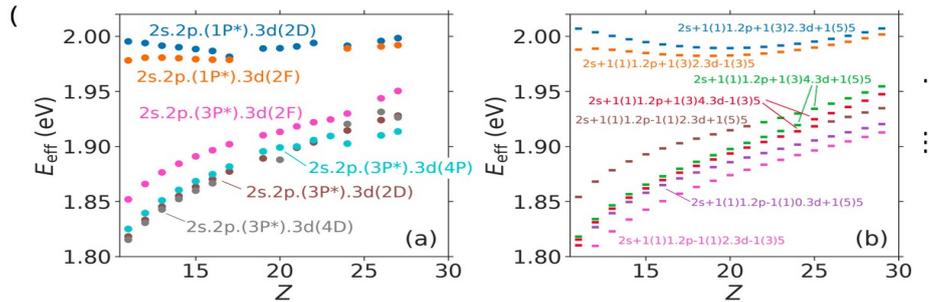


図1 Be様イオン $2s2p3d$ ($J=5/2$) 電子配置に関する等電子系列のエネルギー。縦軸は $E_{\text{eff}} = E / (Z - n_{\text{ele}} + \Delta)^2$ で定義される実効エネルギー。ここで、 E は励起エネルギー、 Z は原子番号、 $\Delta=3$ は遮蔽係数である。左図がNIST-ASDに掲載されている実験データ、右図がFlexible Atomic Codeによるシミュレーション結果である。各準位には便宜的に名前（例えば $2s.2p.(1P^*).3d(2D)$ など）がつけられているが、これが系列中に入れ替わっていることがわかる。

例として図1 に Be様イオン $2s2p3d$ ($J=5/2$) 電子配置に関する等電子系列のエネルギーを示す。この配置の中には複数の微細構造があり、それぞれの微細構造ごとに等電子系列を形成していることがわかる。しかし実験結果（図1(a)）やシミュレーション結果（図1(b)）両者で、微細構造の系列の上下関係が入れ替わる点が存在することがわかる。さらに実験データには欠損している準位もあり、単純に連続性を仮定することは難しい。

本研究では、このような連続関係を自動的に見つけるアルゴリズムを開発した。実験データ (R) と機械学習による予測値 (S) の関係から、最寄りの距離の合計を最小にする組み合わせを探し出す。

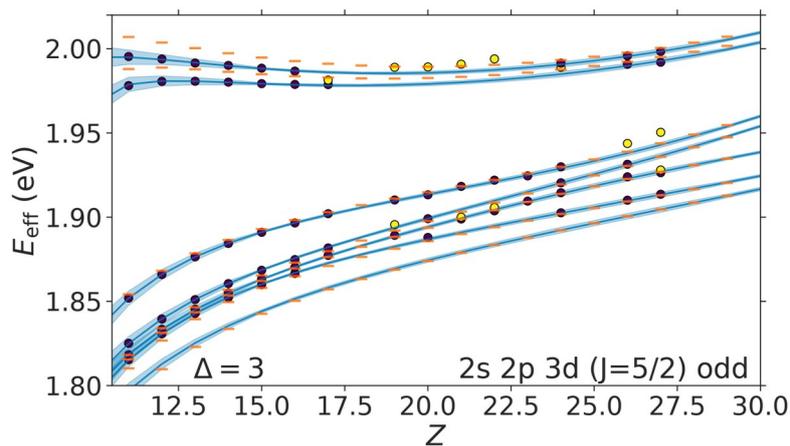


図2 NIST-ASDに掲載されているBe様イオン $2s2p3d$ ($J=5/2$) 電子配置のエネルギー準位（丸印）、FACによるシミュレーション結果（横線）と機械学習による推定結果（実線）。黄色で示した実験データは予測値から有意に外れており、データベースに含まれる誤データであると推測できる。

図2は、NIST-ASDに含まれているBe様イオン $2s2p3d$ ($J=5/2$) 電子配置のエネルギー準位 (丸印)、第一原理シミュレーションによる計算値 (水平線)、シミュレーション値を実験データに近づけることにより求めた推定値 (実線) を示す。黄色丸印で示した実験データは推定値から有意にはずれている。データベースに含まれる誤データであると推測できる。

このような解析を繰り返した結果、おおよそ5%の確率でデータに統計的には説明できない外れ値があることが明らかになった。そのなかには、過去に行われたデータ評価におけるミスが、30年間以上残っているものもあった。さらに、外れ値があることをみこんだ補完により、データベースに載っていない欠損値の推定も可能になった。

B) 未同定の発光線から、新たなエネルギー準位を発見する方法論の開発

上記のA)では、すでに同定されているエネルギー準位がどの程度確からしいかを推定するものであった。このようなエネルギー準位は通常、観測される発光線の波長がエネルギー間隔に対応することを利用して求められる。しかし、大量に観測される発光線を矛盾なく説明するエネルギー準位を求めることは難しい。これまでは経験豊富な研究者の手作業に頼ってきた状況であるそこで本研究では、大量に観測される未同定の発光線波長と、既知のエネルギー準位から新しいエネルギー準位を同定する方法の構築を行った。

具体的には、既知のエネルギー準位セット $E = \{E_1, E_2, \dots, E_n\}$ と発光線の観測により得られたエネルギー差でいまだ同定されていないもの $\Lambda = \{\lambda_1, \lambda_2, \dots, \lambda_m\}$ から、発光線を同定し未知のエネルギー準位を見つけることを考える。実際には単一の準位ペアの差が観測されるエネルギー差として得られるが、本研究では全ての組み合わせを列挙する。具体的には、下準位を E_i としたときの準位の候補は $E_{ij} = E_i + \lambda_j$ ($i = 0, \dots, n, j = 0, \dots, m$) のようにあらわされ $n \times m$ 通りを挙げることができる。なお、選択則を考慮することによってもう少し制限できる。列挙した候補のほとんどは誤った組み合わせであるため、それらにより推定した準位のエネルギーはランダムに分布する一方で、上準位が実在する真の組み合わせの場合は、複数の下準位への遷移が観測されるため候補が集中するはずである。

図3(a)に、NIST-ASDに収録されている中性タングステン原子のエネルギー準位を示す。なお、NIST-ASDには465個の既知のエネルギー準位が収録されている。さらにNIST-ASDには中性タングステン原子の未同定の発光線が1197本収録されている。図3(b)、(c)に、これらの既知の準位と未同定の発光線の全組み合わせから求めたエネルギー準位候補の数密度を示す。数密度分布は有限のノイズフロアに加えて、複数のピーク構造が確認できる。これらのピークを外れ値検出手法に基づき検出することで、新しいエネルギー準位を求めた。このようにして、中性タングステン原子に関して10個の新しい未報告のエネルギー準位を決定した。

本手法は、既知のエネルギー準位と未同定の発光線のデータさえあれば自動的に実行できる。同様の手法を、NIST-ASDに掲載されている他の中性原子のデータに対して適用することで、シリコンや鉄などについて合計20個程度の未報告のエネルギー準位を同定することが出来た。さらに本手法は原子核のエネルギー準位同定にも適応できる。実際、National Nuclear Data Center の管理するENSDF データベースから様々な原子核に対する既知のエネルギー準位と、未同定のガンマ線エネルギーをダウンロードし、本手法を用いることで、合計20個程度のみ報告のエネルギー準位を求めることが出来た。

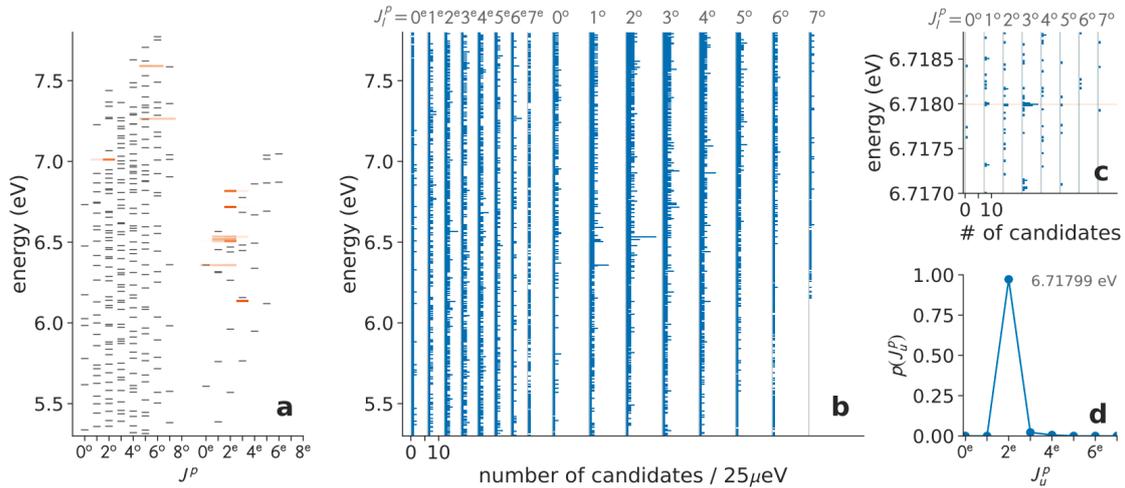


図3 (a) NIST-ASDに収録されている中性タングステン原子のエネルギー準位。(b) 既知のエネルギー準位と未同定の発光線エネルギーから求めた、エネルギー準位の候補の数密度分布。誤った組み合わせから求めたエネルギー準位はランダムに分布する一方、真の準位が存在するエネルギーには多くの候補が集中するため、ピークとなって観測される。(c) bの拡大図。(d) 選択則を検討することで、新たに見つけた準位のパリティ・角運動量量子数の推定も可能である。例えば6.71799 eVのエネルギー準位は高い確率で $J=2$, evenパリティのものであると推測できる。新たに発見した準位を(a)にオレンジ線で示す。

C) 少量の原子データのみから多電子原子発光線の挙動を推定するモデルの構築

A), B)は、大量の原子データを生成・評価する方法論の開発であった。一方で、本テーマは大量のエネルギー準位・遷移が存在する複雑な原子の挙動を確率論を用いて少量のパラメータのみで表現する方法について検討した。

中でも、これまで知られてきた多電子原子スペクトルの性質の1つに、発光線強度のべき乗分布がある。大量に観測される多電子原子からの発光線スペクトルを、それらの強度ごとに数分布を作成するとその分布が強度のべき乗になる、というものである。この経験則が40年前に報告され、他の実験や第一原理計算により再現されてきたものの、その原因は明らかになっていない。本研究では、多電子原子の構造を漁師カオス理論により確率的にモデル化することで、本性質を初めて理論的に説明した。

具体的には、まず多電子原子のエネルギー準位の数密度（準位密度、 $\rho(E)$ ）を考える。これは電子がパウリの排他則に従うことやその組み合わせ則から、 $\rho(E) \propto \exp(E/\epsilon_0)$ というように励起エネルギー E に関して指数関数的に増加することが知られている。 ϵ_0 は増加のエネルギースケールである。

発光強度 I は上準位の占有密度に比例する。高密度プラズマ中では、占有密度分布がボルツマン分布に近づくことが知られている。

$$n(E) \propto \exp(-E/kT_e)$$

上記2つのモデルを組み合わせることで、発光強度 I をもつ発光線の数 $\rho(I)$

$$\rho(I) \propto I^{2kT_e/\epsilon_0 - 1}$$

と表されることがわかった。実際、第一原理計算でもこのべき乗分布やその指数の電子温度依存性が再現された。

5. 主な発表論文等

〔雑誌論文〕 計3件（うち査読付論文 3件/うち国際共著 3件/うちオープンアクセス 3件）

1. 著者名 Fujii Keisuke, Berengut Julian C.	4. 巻 126
2. 論文標題 Power-Law Intensity Distribution of γ -Decay Cascades: Nuclear Structure as a Scale-Free Random Network	5. 発行年 2021年
3. 雑誌名 Physical Review Letters	6. 最初と最後の頁 102502
掲載論文のDOI（デジタルオブジェクト識別子） 10.1103/PhysRevLett.126.102502	査読の有無 有
オープンアクセス オープンアクセスとしている（また、その予定である）	国際共著 該当する

1. 著者名 Fujii Keisuke, Berengut Julian C.	4. 巻 124
2. 論文標題 Simple Explanation for the Observed Power Law Distribution of Line Intensity in Complex Many-Electron Atoms	5. 発行年 2020年
3. 雑誌名 Physical Review Letters	6. 最初と最後の頁 185002
掲載論文のDOI（デジタルオブジェクト識別子） 10.1103/physrevlett.124.185002	査読の有無 有
オープンアクセス オープンアクセスとしている（また、その予定である）	国際共著 該当する

1. 著者名 Nishio Akira, Berengut Julian C., Hasuo Masahiro, Fujii Keisuke	4. 巻 102
2. 論文標題 Population kinetics of many-electron atoms in ionizing plasmas studied using a continuous collisional radiative model	5. 発行年 2020年
3. 雑誌名 Physical Review E	6. 最初と最後の頁 53211
掲載論文のDOI（デジタルオブジェクト識別子） 10.1103/PhysRevE.102.053211	査読の有無 有
オープンアクセス オープンアクセスとしている（また、その予定である）	国際共著 該当する

〔学会発表〕 計6件（うち招待講演 0件/うち国際学会 0件）

1. 発表者名 西尾 顕, Jilian C. Berengut, 蓮尾 昌裕, 藤井 恵介
2. 発表標題 連続衝突ふく射モデルを用いたイオン化進行プラズマ中 多電子原子イオンのポピュレーション分布の研究
3. 学会等名 2020年日本物理学会秋季大会
4. 発表年 2020年

1. 発表者名 藤井 恵介, Jilian C. Berengut,
2. 発表標題 ガンマ崩壊カスケードにおけるべき乗の強度分布
3. 学会等名 2020年日本物理学会秋季大会
4. 発表年 2020年

1. 発表者名 藤井 恵介
2. 発表標題 多電子原子イオンのポピュレーションダイナミクスに対する 縮約モデリングと発光線スペクトルの統計則
3. 学会等名 プラズマ・核融合学会第37回年会
4. 発表年 2020年

1. 発表者名 本宮和明, 蓮尾昌裕, 藤井恵介
2. 発表標題 多電子原子イオンを始めとする多体量子系に対する エネルギー準位同定アルゴリズムの開発
3. 学会等名 プラズマ・核融合学会第37回年会
4. 発表年 2020年

1. 発表者名 藤井 恵介, Jilian C. Berengut,
2. 発表標題 多体量子系における放射崩壊カスケードと スケールフリーランダムネットワーク
3. 学会等名 第76回年次大会
4. 発表年 2021年

1. 発表者名 Keisuke Fujii, Jose R. Crespo Lopez-Urrutia
2. 発表標題 Statistical completion and validation of the NIST Atomic Spectral Database
3. 学会等名 DPG Spring Meetings 2019
4. 発表年 2019年

〔図書〕 計0件

〔産業財産権〕

〔その他〕

<p>多電子原子の発光線数は強度のべき乗に従う（日本語） https://fujiiisoup.github.io/keisukefujii/2021-02-25/Learners-law 原子核はスケールフリーネットワーク構造を示す（日本語） https://fujiiisoup.github.io/keisukefujii/2021-03-13/nuclear-network</p>

6. 研究組織		
氏名 (ローマ字氏名) (研究者番号)	所属研究機関・部局・職 (機関番号)	備考

7. 科研費を使用して開催した国際研究集会

〔国際研究集会〕 計0件

8. 本研究に関連して実施した国際共同研究の実施状況

共同研究相手国	相手方研究機関
---------	---------