

研究論文

セグメント回帰を用いたべき乗則で解釈される スペクトルの前処理法

柳生 進二郎,^{1,*} 吉武 道子,¹ 長田 貴弘¹

¹ 国立研究開発法人 物質・材料研究機構 〒305-0044 茨城県つくば市並木 1-1

*YAGYU.Shinjiro@nims.go.jp

(2024年2月22日受付; 2024年4月25日掲載決定)

セグメント回帰を用いて、べき乗則で解釈されるスペクトルの解析領域を抽出するための自動前処理法について報告する。スペクトルに対してセグメント回帰を適用することで、スペクトルの変化点やその前後の傾きの特徴を得ることができる。これらの特徴を評価・判定することで、解析領域を自動的に抽出することができる。べき乗則で解釈される光電子収量分光スペクトルについて本法を適用した解析例を示す。

Preprocessing Method for Spectra Interpreted by the Power Law Using Segment Regression

Shinjiro Yagyū^{1,*}, Michiko Yoshitake¹, and Takahiro Nagata¹

¹ National Institute for Materials Science, Tsukuba, Ibaraki 305-0044

*YAGYU.Shinjiro@nims.go.jp

(Received: February 22, 2024; Accepted: April 25, 2024)

We report an automated preprocessing method for extracting analysis regions from measured spectral data interpreted by the power law using segment regression. By applying segment regression to spectra, it is possible to obtain the characteristics of the change points and the slopes before and after the change points in the spectrum. By evaluating and judging these features, the analysis region can be automatically extracted. An example of analysis applying this method to a photoelectron yield spectrum with a power law characteristic is shown.

1. はじめに

近年、マテリアルインフォマティクス (MI) の進展により大量のデータが必要になった。また、ロボット技術の発展により実験の自動化[1]が行われ、大量の測定スペクトルが生成されるようになった。これにより、大量のスペクトルを自動で解析処理する方法が望まれている。スペクトル解析では、解析対象領域を特定し(関心領域 (ROI: Region of Interest))、抽出し、その領域に対して解析方法を適用する。ROIの抽出では、スペクトルの全体的な特徴を把握する必要がある。一般的にスペクトルの特徴を把握する方法には、微分、統計量(最大値, 最小値, 平均値, 分散, 中央値, ピーク値など), フーリエ変換などがある。このうち微分(傾き)が求まると、スペク

トル値の変化・特徴を捉えることができる。実際のスペクトルでは、ノイズにより数値的な微分(計算上は2点間の差分を求めている)では、特徴を捉えるような量を求めることは難しい。このため一般的には、スムージング(移動平均)によりノイズ成分を減らして数値微分を行うが、スムージングの程度が問題になる。一方で、ノイズが乗っているデータの特徴を捉える方法として、セグメント回帰がある。セグメント回帰は、複数の回帰直線にてデータをモデル化することができる。この手法を用いることでノイズが乗っているスペクトルに対しても変化点とその前後での傾きを推定することができる。変化点や傾きを評価・判定することで、スペクトルの全体をとらえ、解析に用いる領域 (ROI) を抽出すること

ができる。セグメント回帰と評価・判定のアルゴリズムによって、コンピューターでの自動解析が可能になるとともに、これまで行われてきた人に依存した解析を排除することができる。本研究では、最初にシミュレーションデータに対してセグメント回帰によってROIを切り出すスペクトル解析の前処理の自動化方法について報告する。そして、イオン化ポテンシャルを求める光電子収量分光スペクトル (Photoelectron Yield Spectroscopy: PYS スペクトル) について本法を適用し、ROIを切り出し、その後、我々が以前報告したべき乗解析方法[2]を適用した結果について示す。

2. セグメント回帰と傾き評価

2.1 セグメント回帰

セグメント回帰 (又は *Piecewise Regression* と呼ばれる) は、データセット内の関係性が1つの直線では表現できず、異なる区間 (セグメント) ごとに傾きが異なる直線関係がある場合に使用される統計的手法である。Fig. 1 は、横軸 $x=7$ 付近で傾きが変わるガウスノイズが乗っている計算データに対して、

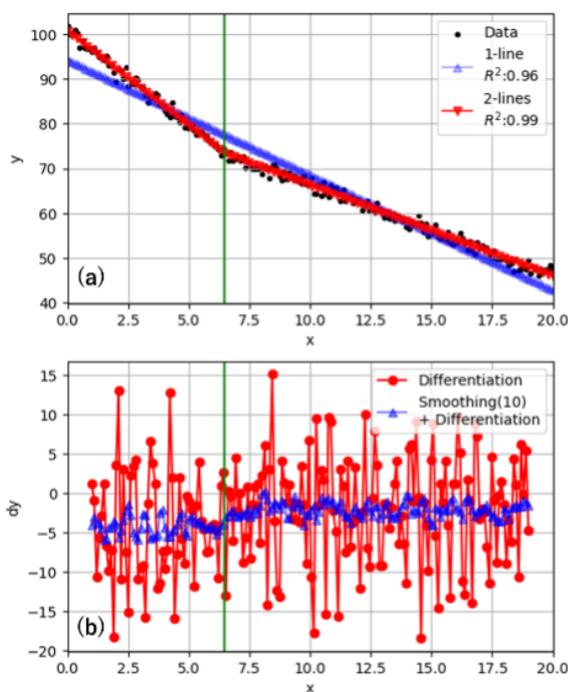


Fig. 1. A comparison of the feature extraction from the data (data with noise whose abscissa tends to change trend around $x=7$) using piecewise regression and numerical differentiation. (a) shows the results of fitting the data with one and two straight lines. (b) shows the results of numerical differentiation with and without the moving average procedure. (color online)

セグメント回帰及び数値微分を適用した結果である。Fig. 1 (a)は、データに対して1つの直線及び2つの直線 (セグメント回帰) による Fitting 結果である。1つの直線による Fitting (∇ マーク) では当てはまりが悪い (当てはまりの評価指標である決定係数 (R^2) が小さい)。 $x=7$ 付近を境に2つの直線で Fitting (Δ マーク) を行うと当てはまりが良いことがわかる (R^2 がより1に近い)。この時の2つの直線の傾きを “微分係数” (厳密には異なる) とみなすことができる。Fig. 1 (b)は、データに対し数値微分を行った結果である。スペクトルに対して、そのまま数値微分を行うとノイズの影響で特徴を把握することができない。特に、 $x=7$ 付近の変化点を検出することができない。スムージング (移動平均) をかけて数値微分を行うと、かろうじて変化点を確認することができる。スムージングの程度は、データ点数やノイズの程度によりスペクトルを見てトライアンドエラーにより決定しなければならない。一方で、セグメント回帰を用いることで、変化点とその前後の傾きの情報が得られる。(今後、傾きの異なる直線が交わる点を、変化点又は、区分位置 (Breakpoint, BP) と呼ぶ。) セグメント回帰のアルゴリズムは以下の通りである。区分数 (変化点の数) を1として与えると、ランダムまたは、与えられた初期値により初期の区分位置を決める。そして、その区分位置の前後で線形回帰を行い、各回帰直線の残差 (データと回帰線との差) を計算する。区分位置を変えて同様の回帰計算を行い、残差が一番小さくなる区分位置を求める。この方法では、区分数が多いほど当てはまりは良くなる (残差は小さくなる)。そのため区分数の設定が、解析者の勘・コツ・経験に左右されてしまうことになるが、情報量基準を導入することで、自動で区分数を推定することができる。情報量基準とは、統計モデルの選択や評価に用いられる指標で、モデルの当てはまりの良さ (残差) とモデルの複雑さ (区分数) のバランスを考慮するために用いられる。データに対して区分数を増やしていき、区分数と当てはまりの程度を情報量基準にて計算し、情報量基準の計算値が一番小さい時の区分数を、適切な区分数とする。セグメント回帰のプログラムの実装は、R[3]や Python[4,5]で行われており公開されている。本解析では、すでに公開されている Python モジュール (Piecewise regression, ベイズ情報量基準 (BIC) による、区分数と当てはまり評価も実装されている) [4] を用いて、BIC によって評価した最適な区分数と区分位置 (BP) 及び区分間の

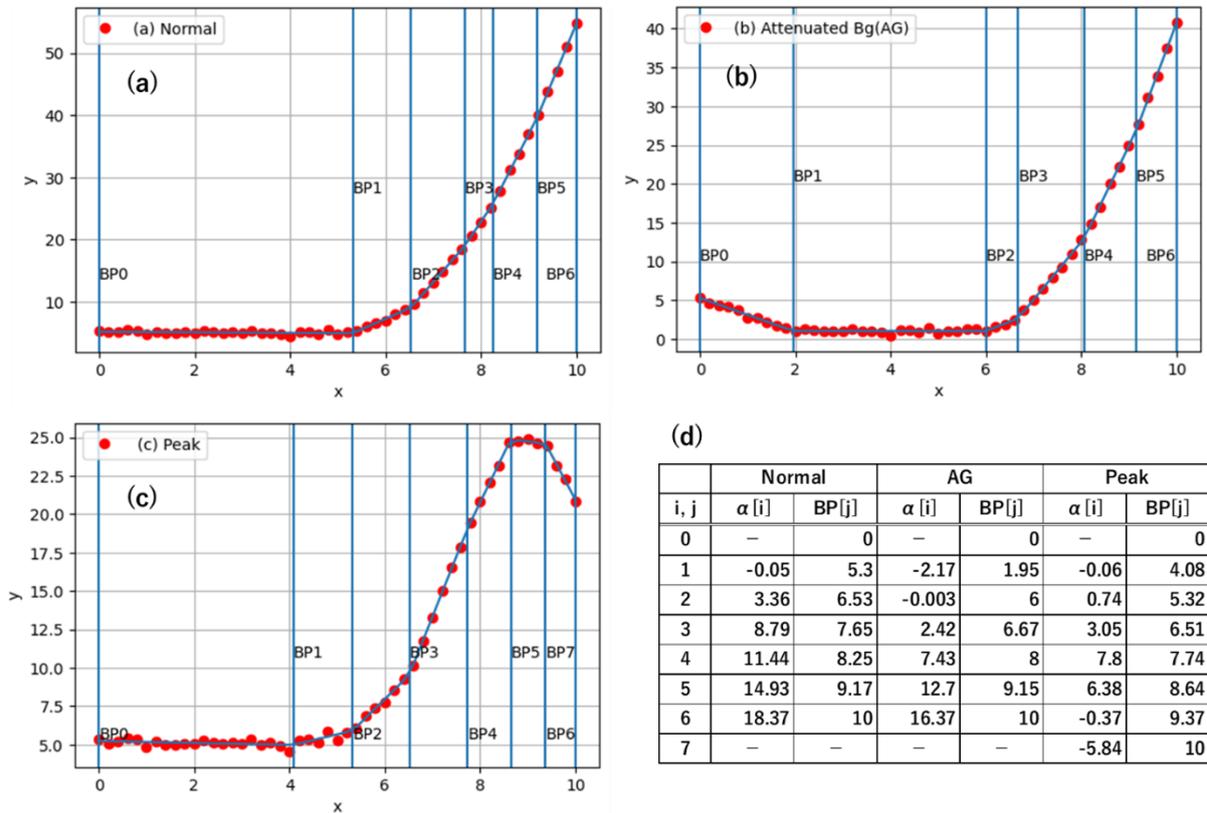


Fig. 2. Results of applying the PW method to a typical power-law calculation spectrum containing Gaussian noise and a power-law calculation spectrum containing components commonly found in experiments, where the line perpendicular to the x-axis is the position of each segment (marked with BP). (a) Normal is the power-law spectrum with constant background on it (the power is set to 2 in the calculation). (b) Attenuated Bg (AG) is a spectrum where the intensity decreases once when x is small and then increases by a power after a constant background. (c) Peak is a spectrum that increases by a power and then decreases in intensity and peaks. (d) shows the BPs and α for the three cases. (color online).

傾き (α) を求めた (今後, 区分を表す英語名 (Piecewise) から, 本方法を PW 法と呼ぶ). ROI を抽出する自動化では, 得られた BP 及び α の値を評価・判断する指標とそのアルゴリズムが必要である.

2.2 区分位置と傾き評価

前述の通りセグメント回帰を用いることで, スペクトルの変化点とその前後の傾きを求めることができる. ガウスノイズを付与した典型的なべき乗計算スペクトルと, 実験で得られるような成分を含むべき乗計算スペクトルに対して, PW 法を適用した結果を Fig. 2 に示す. (なお, 実際にはここで示すスペクトル形状以外に様々あるが, PYS スペクトルによくみられる典型的例を示している.) 各グラフの x 軸に垂直なラインは, それぞれの区分位置 (BP) である. Fig. 2 (a) Normal は定数バックグラウンドが乗っているべき乗スペクトル (計算ではべき乗数を 2 としている.), (b) Attenuated Bg (AG) は x が小さいときに強度が一度減少し, その後定数のバックグ

ラウンドになった後にべき乗で増加するスペクトル, (c) Peak はべき乗で増加したのち, 強度が減少してピークが現れるスペクトル, (d) はそれぞれのスペクトルにおける PW を適用した際の区分位置 (BP) と傾き (α) をまとめた表である. なお, 区分位置の数が各グラフで異なる ((a) 及び (b) では 6 個, (c) では 7 個) のは, BIC による情報量基準により最適な区分数が自動で選ばれた結果である. 3 つのスペクトルともに, べき乗増加領域で BP の間隔が狭くなっていることがわかる. 計算において, べき乗数を 2 としているために, その微分は x の 1 乗であり, x の値に応じて傾きの値は大きくなる. (b) の AG は高感度の電流測定などの場合で, 測定開始とともに電流計にたまっていたチャージが減衰して一定のバックグラウンドになるまで待ち, その後計測する場合などに現れる. (c) の Peak は, 対象としている現象に他の成分が重畳しているときに現れる. 今回は, べき乗増加が始まる閾値とべき乗数が解析の特微量となることから, (a) のようなスペクトル形状の領域の抽出がターゲットになる. 従って ROI の抽出に

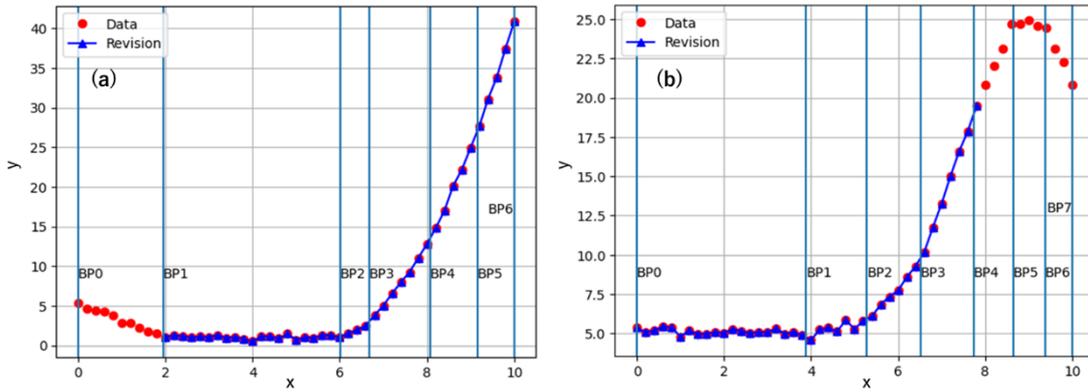


Fig. 3. Results of applying the evaluation algorithm to Fig. 2(b) and (c).

(a) is the result of removing the initial intensity decrease (removed up to BP[1]) and (b) is the result of removing the peak (removed BP[4] and onwards). This decision algorithm can be used to obtain a ROI with AG, Peak, etc. removed.s. (color online).

は、AG のスペクトルでは前半の強度の減少領域 (BP1 まで) を外す必要があり、Peak のスペクトルではピーク付近 (傾きが鈍化する BP4) 以降を外す必要がある。

2.3 判断指標とアルゴリズムとその適用

Fig. 2 で示したスペクトルに対して ROI としてべき乗則に従うスペクトル領域を自動で抽出するには、以下のようなアルゴリズムが考えられる。

- (1) スペクトルに対して PW 法を適用することにより最適な数の区分数及び区分位置 (BP) 及び傾き (α) を計算する。(ここで、各区分の傾き配列 $\alpha[i]$ (区分数を n とすると i は 1 から $n+1$ までの $n+1$ 個), 区分位置 $BP[j]$ (区分数を n とすると j は 0 から $n+1$ までの $n+2$ 個, $BP[0]$ は x の最小値, $BP[n+1]$ は x の最大値である) とする。)
- (2) AG における初期の強度減少領域の削除は、傾きが負になることから、配列の最初から評価して

いき $\alpha[i] < 0$ の領域は、AG 領域と判定することができる。そして領域 $BP[i]$ までを削除する。

(3) Peak の除去は、ピーク近くなると、これまでの傾きに比べて傾きが小さくなることから、配列の最初から比較していき、 $\alpha[i+1] < \alpha[i]$ の時の $BP[i]$ 以降はピーク領域と判定し削除する。

例えば Fig. 2(d) より、AG のスペクトルでは、 $\alpha[1]$ が負になっているのがわかる ($\alpha[1]$ は $BP[0]$ から $BP[1]$ までの区間の傾きである)。Normal や Peak でも Fitting の結果として最初に小さな値の負になっているが、これについては負の値の判定基準を 0 以下ではなく -0.1 以下として無視する。(なお今回は、判定基準値を手動で設定しているが、例えば、ノイズレベルにも依存するがグラフの最大値の 1% の傾き変化を負の値の判定基準とすることもできる)。Peak のスペクトルでは、 $\alpha[5]$ の値が $\alpha[4]$ の値に比べて小さくなっているため、 $\alpha[4]$ までを採用し、 $BP[4]$ までを解析対象として、それ以降を取り除く。Fig. 2(b) と (c) に判定アルゴリズムを適用した結果を Fig. 3 に示す。Fig. 3(a) は、初期の強度減少が取り除かれた結果 ($BP[1]$ まで削除) で、(b) はピークが取り除かれた結果 ($BP[4]$ 以降を削除) である。この判定アルゴリズムを用いることで、AG や Peak などを取り除いた、ROI を得ることができる。なお、実際のスペクトルにおいて、2.1 で仮定したスペクトル以外の形状も現れる場合には、傾きの情報を利用した追加のアルゴリズムも必要になる。

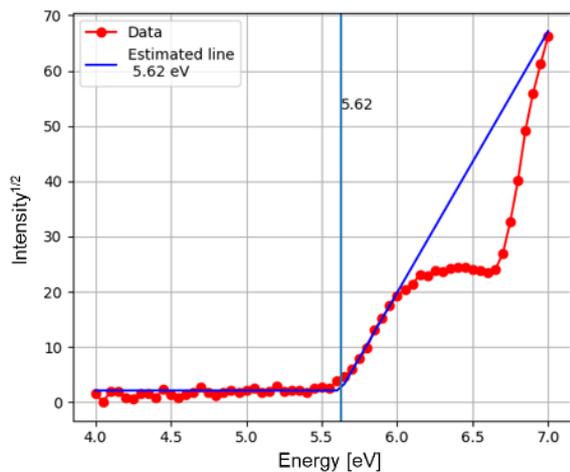


Fig. 4. PYS spectrum in ferrocene powder and the results of the manual analysis. The manual analysis resulted in a threshold value of 5.62 eV. (color online)

3. 実データによる解析 (PYS スペクトルにおける ROI 抽出と閾値推定例)

PYS は、半導体材料におけるイオン化ポテンシャルや金属における仕事関数を測定する方法である。PYS は試料に照射する光のエネルギーを徐々に増加

させ、ある閾値エネルギーを超えると、放出される光電子数がべき乗則に従って増加する。また、電子状態密度に分布があると、スペクトルにピークなどが現れる。粉状のフェロセン ($\text{Fe}(\text{C}_5\text{H}_5)_2$, Aldrich 社製 (F408)98%, 有機半導体のドナー材料) における PYS スペクトル (縦軸は PYS 強度の 1/2 乗した値, 横軸はエネルギーである。) [5] とその手動解析結果を Fig. 4 に示す。スペクトルでは、ピークが現れその後、再度強度が増加している。PYS による閾値解析では、バックグラウンドから最初に立ち上がる位置を閾値として求める。手動解析の結果、閾値は 5.62 eV であった。(手動解析では、PYS 強度の 1/2

乗を取り、バックグラウンドと 1/2 乗で直線化された傾きとの交点から閾値を求める。) 自動解析を行うには、ピーク (およそ 6.2 eV) より小さいエネルギー領域を解析領域として抽出する必要がある。PW 法を用いてスペクトルを評価し抽出した結果を Fig. 5(a) に示し、その時の BP 及び α を (b) に示す。グラフの縦軸は PYS 強度で、横軸はエネルギーである。Fig. 5(b) より、 $\alpha[4]$ が $\alpha[3]$ の値よりも小さくなっていることから、BP[3] まで解析に利用すると判断できる。抽出したスペクトルに対して、我々が開発したべき乗数と閾値の自動推定方法 [2] の 1/n-Power Plot 法を用いて解析した結果を Fig. 6 に示す。有機半導

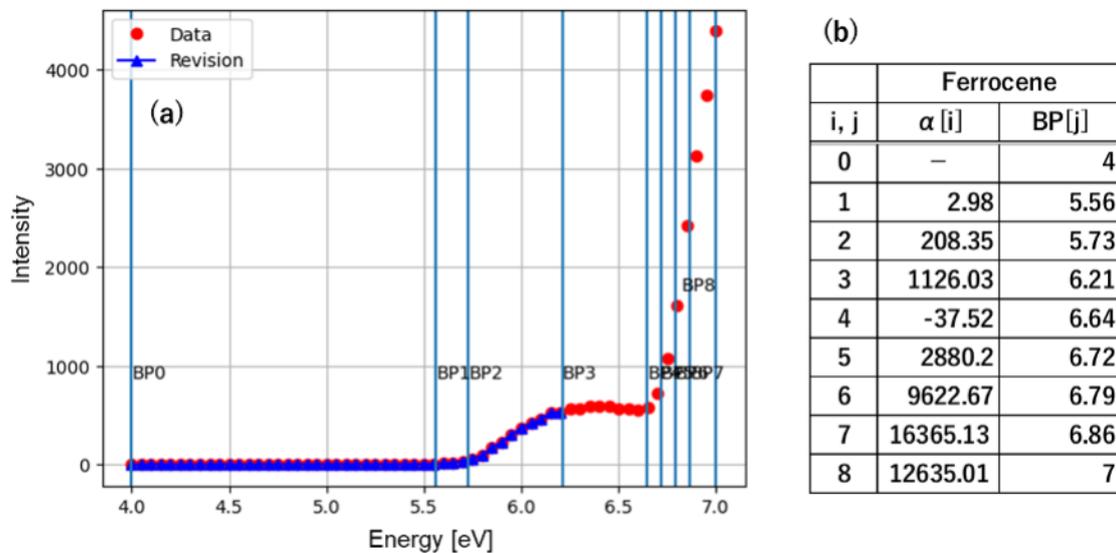


Fig. 5. Results of evaluating the spectrum using the PW method and extracting the analyzed region. The vertical axis of the graph is the PYS intensity and the horizontal axis is the energy. For automatic analysis, the energy region smaller than the peak (approximately 6.2 eV) should be extracted as the analysis region. (b) shows the BP and α when the PW method is applied to the PYS spectrum. (color online).

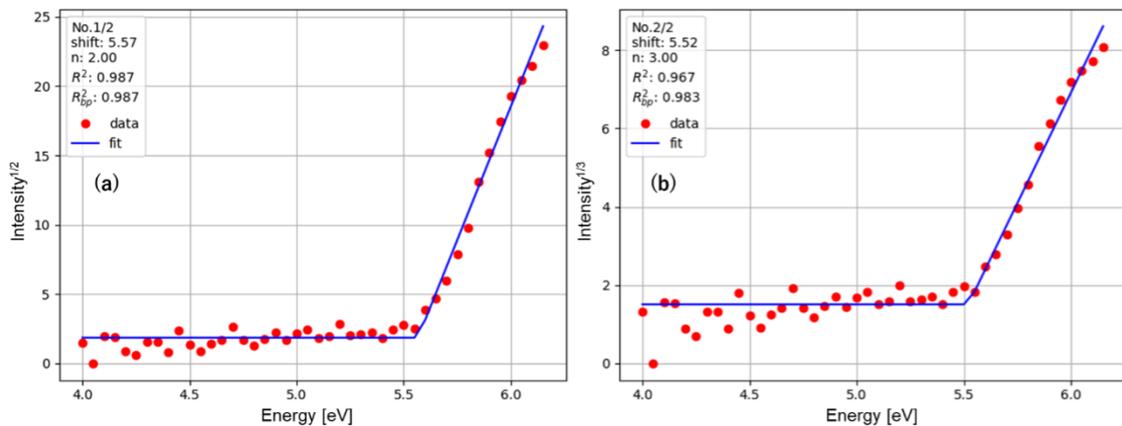


Fig. 6: Results of the analysis using the 1/n-Power Plot method, which is an automatic estimation method for power multipliers and thresholds, for the spectra extracted in Fig. 5. In the case of organic semiconductors, the multiplier (n) was evaluated with the corresponding power, as it could be 2 or 3. (a) and (b) show the results of the fitting with the 1/2 power applied to the PYS intensity and the 1/3 power applied to the PYS intensity, respectively. Comparing the coefficient of determination, the fit was better when 1/n=1/2, with a threshold value of 5.57 eV. (color online)

体材料の解析べき乗数(n)として2と3の可能性があるので、それぞれのべき乗数で評価した。Fig. 6 (a)はPYS強度に1/2乗を適用し、(b)はPYS強度に1/3乗を適用し、それぞれFittingを行った結果である。(グラフ中のShiftは閾値、 n はべき乗数、 R^2 はデータ全体の決定係数、 R^2_{bp} は閾値以降の決定係数である。)決定係数を比較すると $1/n=1/2$ の時の方が1に近く、その時の閾値は5.57eVであった。手動解析との閾値の差は0.05eVであり、測定間隔(ステップ0.05eV)程度の違いしかないことから、手動解析の値と同等であると考えることができる。また、この解析からべき乗数が1/2乗の方が確からしい結果となった。1/2乗は、金属での解析で得られることから、金属的な性質が高いと考えることもできる。PW法及び、判定アルゴリズムを用いて解析領域を抽出し、その後べき乗における自動推定法を用いることで、PYSスペクトルデータを自動で解析することができる。

4. まとめ

セグメント回帰(PW法)は、データをいくつかの区間に分割し、それぞれの区間に対して異なる直線を適用することで、複数の回帰直線にてデータをモデル化することができる。この手法を用いることでノイズが乗っているスペクトルに対しても変化点とその前後での傾きを推定することができる。

べき乗則で増加する成分以外の実験で見られるような成分(バックグラウンドの減少、べき乗増加の領域でのピーク)が含まれ、ノイズが乗っているシミュレーションデータに対して、PW法を適用した。そして、変化点前後の傾きの値を評価・判定するアルゴリズムを作成し適用することにより、解析領域を自動で抽出することができる。

有機半導体材料のフェロセン粉末で、ピークを持つような光電子収量分光スペクトルについて本法を適用し、自動で解析領域を抽出した。そして、我々がすでに報告しているべき乗における自動解析法を適用することで、自動解析を行った。その結果、手動で解析した値とほぼ同程度の結果が得られた。

PW法については、すでに公開されているPythonパッケージを利用した。

これまで解析者の勘、コツ、経験で決めていた解析領域に対して、この手法を利用することで、統計的に解析領域を特定することができる。これにより、“なぜこの領域を選んだか”との問いに対して、回答を示すことができ、解析の人依存を回避すること

ができる。

5. Suplimental

すでに公開されているPythonパッケージを利用して、本報告の一連の処理を行ったソースコードをサイトに公開予定である。

6. Reference

- [1] R. Shimizu, S. Kobayashi, Y. Watanabe, Y. Ando, T. Hitosugi, *APL Mater* **8**, 111110 (2020).
doi: 10.1063/5.0020370
- [2] S.Yagyu, M.Yoshitake, T.Nagata, *J. Surf. Anal.* **30**, 98 (2023).
doi:10.1384/jsa.30.98
- [3] <https://cran.r-project.org/web/packages/segmented/segmented.pdf>
- [4] <https://github.com/chasmani/piecewise-regression>
- [5] https://github.com/cjekel/piecewise_linear_fit_py
- [6] D. Mahdaoui, C. Hirata, K. Nagaoka, K. Miyazawa, K. Fujii, T. Ando, M. Abderrabba, O. Ito, M. Takagi, T. Ishimoto, M. Tachikawa, S. Yagyu, Y. Liu, Y. Nakajima, Y. Nemoto, K. Tsukagoshi, T. Wakahara, *J. Mater. Chem. C*, **10**, 3770 (2022).
doi.org/10.1039/D1TC05545C

査読コメント, 質疑応答

査読者 1 鈴木 峰晴 (SA コンサルティング)

光電子収量分光スペクトルの自動解析を可能とする提案であり, 大量データの処理には不可欠な内容と判断できるので, JSA への掲載を勧めます. ただ, 読者の誤解を避けるために下記の検討をお願いします.

[著者]

お忙しいところ, 丁寧に査読していただきありがとうございました. 再度検討を行う機会を得たことに感謝いたします. ご指摘を踏まえ下記の通り改訂いたしました.

[査読者 1-1]

本報告は PYS スペクトルの低エネルギーのフラットな領域から指数関数的に上昇する部分を如何に自動的に抽出するかに関するものです. 一方, タイトル「セグメント回帰を用いたスペクトルの前処理法」はかなり一般的な内容を推測させます. JSA の読者にとって「スペクトル」と言われたときには自分の経験に基づき, (a)XPS のような narrow ピーク (上昇から下降) または段差構造上に鋭いピークを有する wide スペクトル, (b)AES のような巨大なバックグラウンド上のピークまたは微分スペクトル, (c)SIMS のようなデルタ関数的なスペクトル, (d)XRD, 光学的分光スペクトルや IR のようなスペクトル, などを思い浮かべるでしょう. ですので, タイトルに「光電子収量分光スペクトル」に関する考察であることを明記するのが良いのではないのでしょうか. もしくは, XPS のようなスペクトルにも適用可能なら付記でも良いので, 事例を示すのが良いでしょう.

[著者]

ご指摘の通り, 光電子収量分光スペクトルのようなべき乗則で解釈するようなスペクトルが主な対象になります. ご指摘を踏まえ, タイトルの対象を明確にするために「べき乗則で解釈される」の修飾語をスペクトルの前に追記し, 「セグメント回帰を用いたべき乗則で解釈されるスペクトルの前処理法」に修正いたしました. それに伴い英語のタイトルも修正しました. 「Preprocessing method for spectra interpreted by the power law using segment regression」

[査読者 1-2]

レポート全体を通じて「変化」「変化点」「change」「change point」という用語が多用されています. 「変化」という用語は, 『何の』『どのような』変化かを明記しないと著者や読者の主観により曖昧が生じるかと思います. レポートの初めの部分でそれを明確に定義して, 『以降本レポートでは「変化」「変化点」「change」「change point」と呼ぶ』, とされるのが良いのではないのでしょうか.

[著者]

変化・変化点について整理し, 以下のように記載しました.

2.1 の本文中, 「今後, 傾きの異なる直線が交わる点を, 変化点, 区分位置 (Breakpoint, BP) と呼ぶ. 」と追記しました.

[査読者 1-3]

セグメント回帰法の最適化 (分割数, 直線フィッティング=傾き決定) に関する評価基準 (指針, ガイドライン) に関して簡単に紹介していただけるとよろしいかと思います. 例えば, Fig. 2(a)はこの結果の optimization は, どのように評価されているのでしょうか.

[著者]

データに対して, セグメント回帰を行う場合, 解析者はパラメータとして区分数を与える必要があります. 情報量基準を用いることで区分数も自動推定することができます. 実際には, 最大区分数を指定し, 区分数 1 から最大区分数まで順番にセグメント回帰を行い, その時のパラメータ数と当てはまりからベイズ情報量基準(BIC)を計算し評価して (BIC が一番小さいもの), 最適な区分数を決定します. (なお同時に, 区分位置およびその前後の傾きも求められます.) 情報量基準は, モデル選択時に使用される統計量で, モデルの適合度を評価します.

具体的には, モデルの観測データへの適合度 (データとモデルとの残差) を示す最大対数尤度と, モデルの複雑さ (パラメータ数) に対するペナルティを組み合わせています. Fig. 2 において, (a)及び(b)では区分数が 6 個に対して, (c)では, 7 個を選び出しています.

2.2 の本部長中に「なお, 区分位置の数が各グラフで異なる ((a)及び(b)では 6 個, (c)では 7 個) のは, BIC による情報量基準により最適な区分数が自動で選び

だされた結果である。」と追記しました。

[査読者 1-4]

「2.2 変化位置と傾き評価」の最後の部分「ROI の抽出には、AG のスペクトルでは前半の強度の減少領域 (BP1 まで) を外す必要があり、Peak のスペクトルではピーク付近 (傾きが鈍化する BP4) 以降を外す必要がある。」は、PYS 解析の領域設定の肝になる部分かと思えます。PYS に不慣れな読者にも、なぜこの領域の解析が必要十分なのかが理解できるように説明を加えていただくのがよろしいかと思えます。

[著者]

まず、指摘事項 1 での指摘からタイトルをべき乗則で従うスペクトルと限定しました。べき乗解析でのパラメータは、閾値とべき乗数であることから、Fig. 2(a)のようなスペクトル形状の領域を抽出することがターゲットになります。

このことを踏まえ、2.2 の本文中に「今回は、べき乗増加が始まる閾値とべき乗数が解析の特徴量となることから、(a)のようなスペクトル形状の領域の抽出がターゲットになる。」と追記いたしました。

[査読者 1-5]

上記にも関連しますが、Fig. 4 のピーク (ショルダー構造) に物理 (物性) 的な意味はあるのでしょうか。本レポートの主題とは外れるかと思いますが、どのようなときに PYS でこのようなスペクトル構造が現れるのか教えていただけますか。

[著者]

ご指摘の通り、Fig. 4 のスペクトルは、閾値以外に多くの情報を含んだものとなります。(なお今回の関心事である閾値の推定には、べき乗で増加する領域のみが解析対象になり、論文で記述の通り、それ以降のデータは、解析対象外としてしまいます。) PYS の原理は、試料に照射するエネルギーを徐々に高くし、その際に放出される光電子をカウントするものです。そのために、エネルギーごとに微分 (数値微分) することで状態密度を得ることができます。Fig. 4 のスペクトルでは途中で増加が止まり、そしてまた増加が始まることから、状態密度が一様増加でなく分布を持っていることがわかります。

また、参考までに、PYS スペクトルとそれを数値微分したグラフを添付いたします。

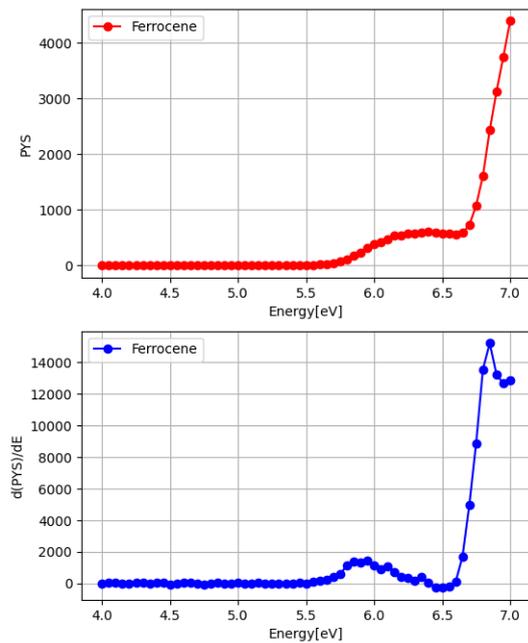


Fig. 6のスペクトルを数値微分した結果 $d(\text{PYS})/dE$ は状態密度になる

査読者 2 井上 雅彦 (摂南大学)

本論文は取得したスペクトルに対して自動解析を適用する関心領域 (ROI) を自動抽出する方法に関するものです。いくつか補足していただいた方が良く思われるところがありましたので、ご検討ください。

[査読者 2-1]

些細なことですがちょっと気になりました。p.4 左カラム上から 2 行目。「PW 法」が本手法の呼び名であるのは良いのですが、なぜ PW (特に W) と名付けられたのかよくわかりません。Python モジュールの名称に由来するのでしょうか? P や W に特に意味がないのでしたら無視してください。

[著者]

ご指摘の通り PW 法は、Python のモジュール名から由来しています。Python のモジュール名は、分布を 1 次関数で表現する区分線形関数 (Piecewise linear function) があり、この名前に由来していると思われます。このことを踏まえ、Piecewise から略称として PW 法としています。

2.1 章の本文中に「(今後、区分を表す英語名 (Piecewise) から、本方法を PW 法と呼ぶ)。」と追記いたしました。

[査読者 2-2]

Fig. 3 を見ると Fig. 2(b)AG と Fig. 2(c)Peak のスペクトルから, Fig. 2(a)Normal と同等の ROI 領域をうまく抽出できており, 提案されている判断指標・アルゴリズムの有効性ははっきり確認できますが, 実験で得られるスペクトルのパターンとしてはこの二つのタイプを考慮するだけで十分なのでしょうか?あるいはこのアルゴリズムではうまくいかない場合も考えられるのでしょうか?もし本手法の適用範囲・制限があるならその辺りの考察・コメントがあると良いのではないかと思います。

[著者]

今回の論文では, PYS スペクトルでよく遭遇するスペクトルの例として示しています。実際には, いろいろなスペクトルがあります。従って, 論文で示した2つのタイプだけで十分かどうかは, 解析対象とするスペクトルによって検証が必要になります。本論文では載せていませんが, バックグラウンドが定数ではなく, 1次関数でその上にべき乗増加が乗っているスペクトルに遭遇することがあります。例えば, 超伝導線材の臨界電流を決める際の IV 測定スペクトルや, バンドギャップを決めるときの吸収スペクトルなどに見られます。この場合, 解析では, 一次関数のバックグラウンドを取り除くまたは推定する処理が必要になります。この推定については, やや複雑であり現在検討中であります。(この推定についても, PW 法を用いて変化点を見つけ, その前後の傾きと判定指標との比較を行うこととなります。判定指標としてスペクトル全体の傾向をからどのような判定指標が必要かを検討しています。)

今回示した手法 (PW 法) は, 変化点およびその前後の領域の傾きを自動で見つける方法です。論文では, 判定指標 (アルゴリズム) と傾き情報の比較からべき乗則に従うスペクトルの領域の抽出例を示しています。実際に自動実験などの系統的な実験における自動処理では, いくつかのサンプルスペクトルから自動化の判定指標を考える必要があります。

2.2 の本文中に「(なお実際には, ここで示すスペクトル形状以外に様々あるが, PYS スペクトルでよくみられる典型的例を示している。)」

2.3 の本文中に「なお, 実際のスペクトルにおいて, 2.1 で仮定したスペクトル以外の形状も現れる場合には, 傾きの情報を利用した追加のアルゴリズムも必要になる。」と追記いたしました。

[査読者 2-3]

p.5 左カラム下から5行目。(なお今回は, …もできる。)の部分。属人的な解析を排除するためにはできればここも自動化が望ましいと思います。最大値の桁数などを参考にしてどのように基準値を設定するのか, 可能ならもう少し具体的に書いていただくと読者の参考になると思います。なお, スペクトルのノイズレベルも関係するように思いますが, いかがでしょうか?

[著者]

例えば, Fig. 2(a)を参考に考えると, 最大強度 (55程度) の1%の変化の傾きを考えると $(55 \times 0.01) / 10$ であり, 傾きは 0.055 となります。これを一つの基準として考えると -0.055 よりも小さい傾きを意味のある傾きと考え, 判定の基準値として設定することができます。これは判定基準の考え方の一つではありますが, 自動化対象とするスペクトルの特徴などを考慮する必要があります。また, ご指摘の通り, ノイズレベルについては非常に関係しますが, ノイズレベルとの関係の考察についてはまだできておりません。

2.3 の本文中には「(なお今回は, 判定基準値を手動で設定しているが, 例えば, ノイズレベルにも依存するがグラフの最大値の1%の傾き変化を負の値の判定基準とするなども可能である。)」と追記いたしました。

編集部注

査読者の指摘箇所 (ページ等) を印刷版の相当位置に修正しました。