

研究論文

べき乗則で解釈されるスペクトルの閾値及び べき乗数の自動解析方法

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

(2023年10月19日受付; 2023年12月4日掲載決定)

べき乗則に従うスペクトルのべき乗数と閾値（べき乗増加の開始位置）の自動推定方法（1/n-Power Plot 法および Shift-Log-Log Plot 法）を提案する。1/n-Power Plot 法ではべき乗数の事前情報がある場合に利用し、べき乗数の逆数をスペクトルに適用することで増加領域を直線化し閾値を求める。Shift-Log-Log Plot 法は事前情報がない場合に利用し、閾値を考慮した両対数プロットにより、べき乗数を推定する。これらの方法を用いて光電子収量スペクトルと電圧 - 電流スペクトルから、閾値とべき乗数の自動推定を行った。

Automatic Analysis Method for the Threshold and Power Exponent of Spectra Interpreted by the Power Law

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

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

* YAGYU.Shinjiro@nims.go.jp

(Received: October 19, 2023; Accepted: December 4, 2023)

Automatic estimation methods (1/n-Power Plot and Shift-Log-Log Plot methods) are proposed to estimate the power exponent of a spectrum following a power law and the threshold at which the power begins to increase. The 1/n-Power Plot method is used when there is prior information on the power exponent and the inverse of the power exponent is applied to the spectrum to linearise the increasing region and obtain the threshold. The shift-log-log plot method is used when there is no a priori information and estimates the power exponent by a log-log plot taking into account the threshold. These methods were used to automatically estimate the threshold and power exponent from the photoelectron emission spectrum and the voltage-current spectrum.

1. はじめに

近年、材料の高機能化と測定装置の自動化により、大量のスペクトルが生成されている[1]。これらのスペクトルから有益な情報や特徴を高速で抽出するスペクトル解析の自動化が重要になってきている。現状の自動解析は光電子分光（XPS）スペクトルなどの一部のスペクトルデータで行われているが[2]、その他のスペクトルでは手動解析のみである。手動解析では多くの時間を要するとともに、解析者による解析の揺らぎが避けられず、定量評価が難しい。

我々はこれまで、半導体材料におけるイオン化ポテンシャルの値を決める光電子収量スペクトル（PYS スペクトル）について、自動解析の検討を行ってきた[3]。PYS スペクトルは、入力変数の値（試料に照射する光のエネルギー）を徐々に上げることで、出力値（放出される光電子数）がべき乗則に従って増加する。類似のべき乗則に従うスペクトルとして、超伝導材料における臨界電流密度とその非線形性能指数(n 値)を決める電圧-電流スペクトル（VI スペクトル）、半導体材料におけるバンドギャップを決め

る透過・反射スペクトル (UV スペクトル) などがある。このようなべき乗則に従うスペクトルでは、べき乗数と、出力値がある入力値を境にべき乗に従って増加が始まる閾値 (原点からの移動量) が特徴パラメータになっている。このパラメータの推定には、理論的なモデルによる事前情報 (べき乗数の値) の有無により 2 つの方法がある。事前情報が与えられている場合には、このべき乗数の逆数をスペクトルに適用することで、測定スペクトルを直線化することができる (今後、 $1/n$ -Power Plot と呼ぶ)。この直線とバックグラウンドの交点を求め、その値が閾値となる。多くの場合この解析は手動で行われている。事前情報がない場合は、べき乗数を推定するために両対数プロット (Log-Log Plot) が利用される。

Log-Log Plot は、横軸と縦軸の両方が対数スケールでプロットされたグラフである。このプロットにより、べき乗則に従って変化するデータは直線で表される。この直線の傾きからべき乗数を求めることができる。実際には、閾値が原点から移動している場合やバックグラウンドの影響で、両対数プロットは 1 次関数の直線にならず、領域によって傾きが異なる。そのために、超伝導におけるべき乗数 (n 値) の推定では解析領域を限定してその領域の傾きを求めている。我々は、閾値やバックグラウンドを考慮に入れた両対数プロットを行うことで、プロットデータが 1 次関数となることを見出した[4]。 (今後このプロットを Shift-Log-Log Plot と呼ぶ。) このプロットを用いると解析領域を選別することなくべき乗数を求めることができる。

べき乗で増加するスペクトルの解析では、手動に

よる解析が行われているのが現状である。手動解析の問題点は、解析者によって異なる解析値が求められてしまうことである。

本論文では、最初にべき乗数の事前情報がある場合の $1/n$ -Power Plot 法における Fitting による自動解析方法について提案を行い、そのシミュレーション結果について述べる。その際、バックグラウンドの閾値への影響についても説明する。次に両対数プロットについて説明した後、我々が提案する原点からの移動量 (閾値) を考慮に入れた Shift-Log-Log Plot 解析の自動アルゴリズムとそのシミュレーション結果について報告する。最後に、実際の PYS スペクトル及び VI スペクトルに、これらの解析法を用いた解析結果を示す。

2. べき乗数の事前情報がある場合の解析方法 ($1/n$ -Power Plot)

測定値 (y) が入力値 (x) に対してべき乗で増加する場合は以下の式で書き表すことができる。

$$y = \begin{cases} d & (x - b < 0) \\ a \cdot (x - b)^c + d & (x - b \geq 0) \end{cases} \quad (1)$$

ここで、 a は定数、 b はシフト (又は shift, 閾値とも本論文では記載している)、 c はべき乗数 (n と記載)、 d はバックグラウンドである。UV スペクトルや PYS スペクトル解析では、べき乗数は理論モデルによって決まった値が用いられている。UV スペクトルからバンドギャップを求める方法として、Tauc-Plot 法[5]がある。Tauc-Plot では、強度 (吸収係数) に光のエネルギーをかけたものに対して、直接遷移ならば 2 乗、間接遷移ならば $1/2$ 乗を適用する

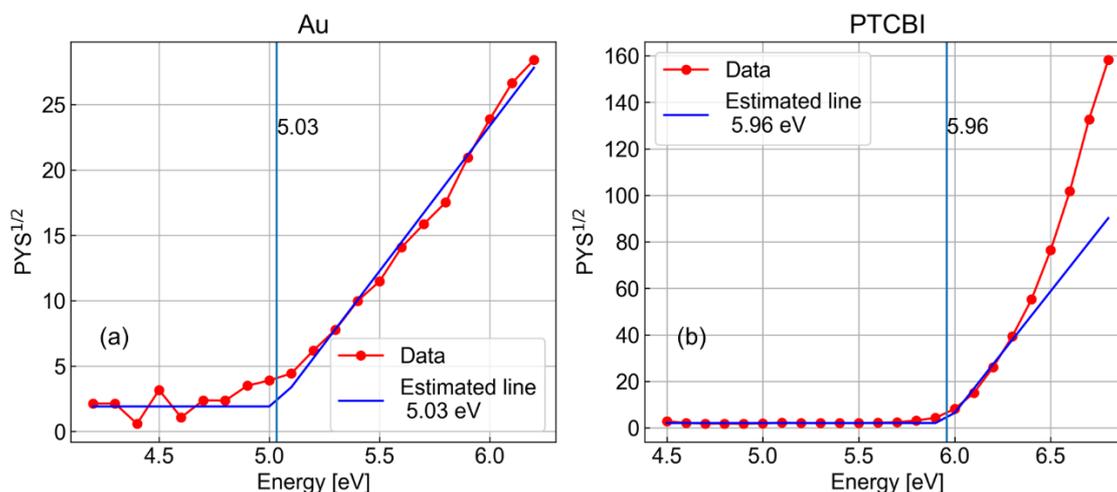


Fig. 1. PYS spectra and their manual analysis (a) Au, (b) organic semiconductor material (PTCBI: 3,4,9,10-Perylenetetracarboxylic Bisbenzimidazole). The vertical axis is the photoelectron yield per photon to the power of $1/2$, the horizontal axis is the irradiation light energy, and the estimated line is the result of manual analysis. (color online)

ことで測定スペクトルを直線化することができる。PYS スペクトルでは、強度（1光子当たりの光電子収量、今後 PYS 強度と記載）が金属では2乗[6]、半導体であれば3乗[7,8]に従って増加すると考えられるために、閾値を求めるために強度の1/2, 1/3乗の適用により直線化できる[9]。この直線ラインとバックグラウンドラインとの交点から閾値を求める。現状、多くの場合手動にて求められている。金属（Au）と有機半導体材料（PTCBI：3,4,9,10-Perylenetetracarboxylic Bisbenzimidazole）におけるPYSのスペクトルとその手動解析例をFig. 1に示す。縦軸はPYS強度を1/2乗したものの、横軸は照射光エネルギーである。グラフにはPYS強度に1/2乗を施したデータ及び手動にて閾値を推定した線が示されている。なお、有機半導体の解析では、1/2, 1/3乗のどちらを用いるべきであるかの議論がある[5]。この解析では1/2乗を用いている。1/2乗を適用したグラフに対して、(a)のAuでは、閾値以降の全領域に渡って直線的に増加している。一方で、(b)のPTCBIでは閾値近くの直線領域に対して線を引き閾値を求めている。エネルギーの高い領域での強度は放物線のように増加しており、1/2乗の適用では直線化が不完全であると考えられる。手動解析の問題は、複数の引き方の候補がある場合、解析者の経験

で直線が引かれ閾値が求められていることである。

2.1 自動閾値推定法

同一スペクトルに対して解析者によって異なる閾値が求められてしまう問題の解決には、数理的に説明できる自動推定法が必要である。そこで、自動推定を行うために以下のようなFitting関数と最小化関数を用いた。Fitting関数として深層学習の活性化関数で用いられるReLU関数[10]を拡張した関数（拡張ReLU関数）

$$\text{exReLU} = \begin{cases} d & (x - b < 0) \\ c'x + d & (x - b \geq 0) \end{cases} \quad (2)$$

を用い、最小化方法として絶対誤差法を用いた。拡張ReLU関数は、閾値(b)まではバックグラウンド(d)で、閾値以降は傾き(c')で増加する関数であり、閾値で微分が不連続となる。実際のスペクトルでの閾値領域では、拡張ReLU関数のようなステップ応答ではなく、様々な要素が組み合わせられるために連続的になだらかに変化する。そのために拡張ReLUと閾値領域では、ずれが大きくなるために、ずれに対してロバストである絶対誤差法を用いる。

2.2 1/n-Power Plot と自動推定のシミュレーション

(1)式に対して(a,b,c,d)=(2,2,3,1)のパラメータを

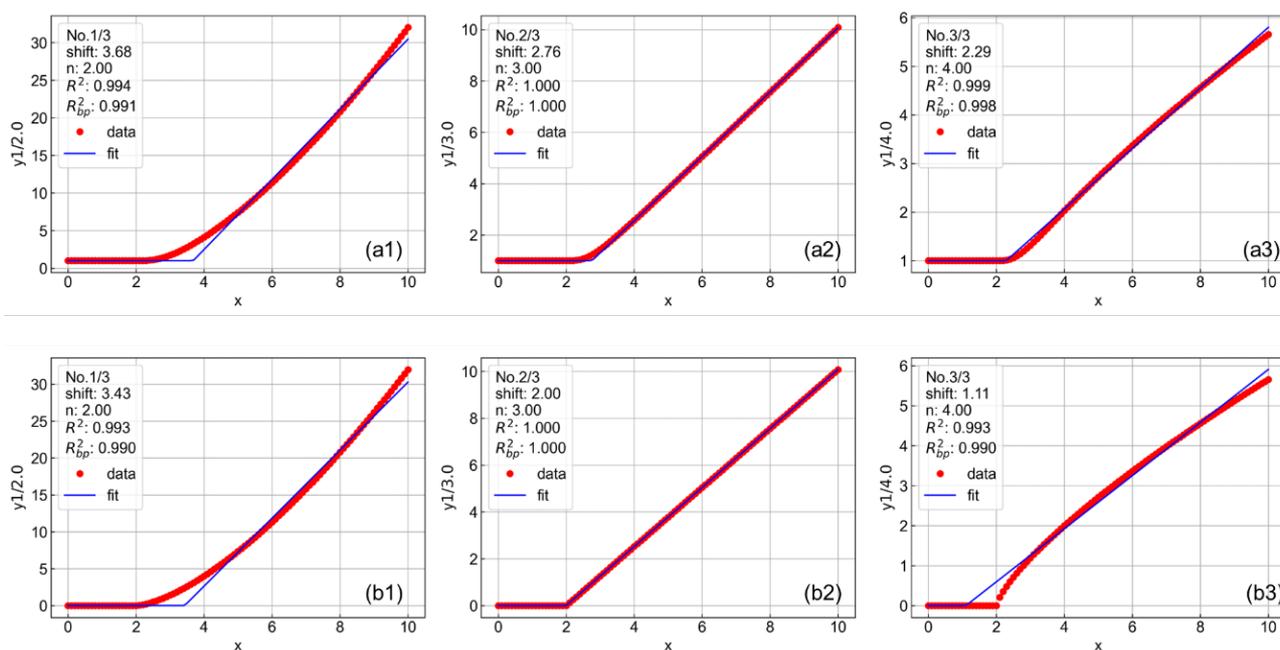


Fig. 2. Fitting results (1/n-Power Plot) for spectra calculated using the parameters (a,b,c,d)=(2,2,3,1) in equation (1) by applying the inverse of several powers (2,3,4). (a1), (a2) and (a3) are cases where the inverse of the power has been applied to the spectrum without background correction, while (b1), (b2) and (b3) are cases where the inverse of the power has been applied after subtraction of the background component (with background correction). R^2 is the coefficient of determination calculated for the entire spectrum coefficient of determination, and R^2_{bp} is the coefficient of determination calculated for regions where x is greater than the threshold value. (color online)

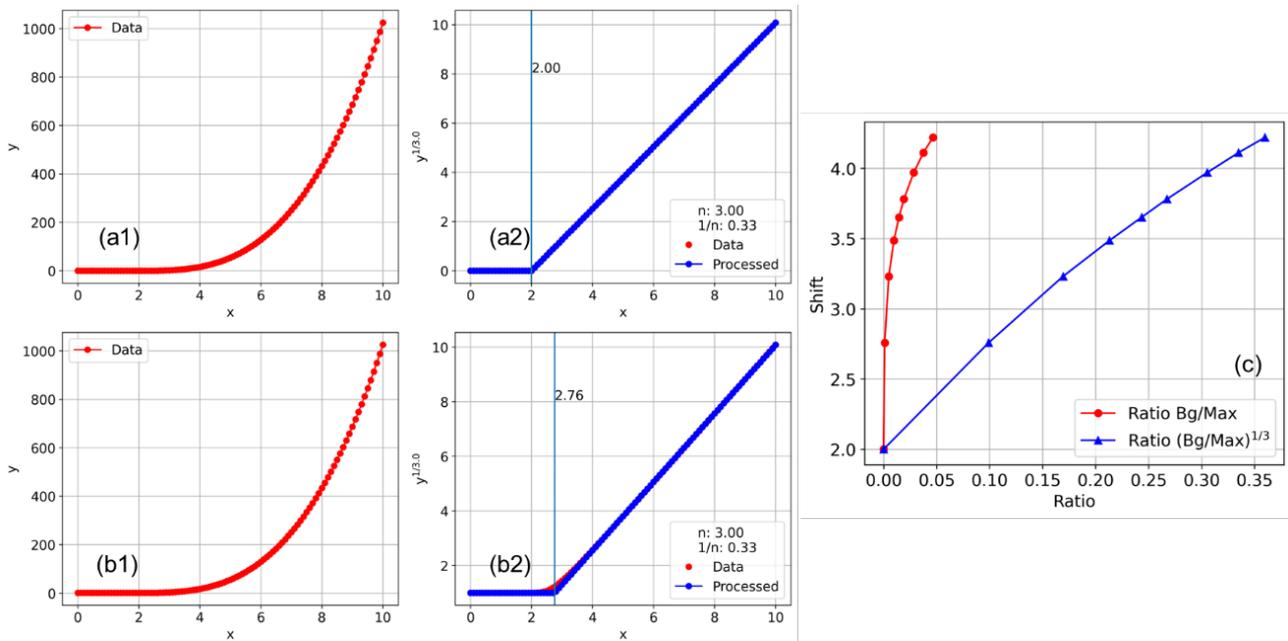


Fig. 3. Variation of shift with different backgrounds. (a1) is the result when the background is 0. (a2) is the result of applying 1/3 power to (a1) for fitting. The shift corresponds to the set value. (b1) is when the background is 1, (b2) is the result of applying 1/3 power to (b1) and fitting results. Although the intensity ratio of the background to the maximum is close to zero, the shift moves to higher values. (c) shows the change in shift and intensity ratio as the background value is varied from 0 to 50. (color online)

用いて計算したシミュレーションスペクトルに、3つのべき乗数の逆数 (1/2, 1/3, 1/4) を適用した結果、Fitting 及びその当てはまり評価(決定係数)を行った結果を Fig.2 に示す。解析に当たり (1) 式にべき乗の逆数をそのまま適用する場合 (バックグラウンド補正なし) と、以下の (3) 式のようにバックグラウンド成分の d を左辺に持って行ってから、べき乗数の逆数を両辺に適用する場合 (バックグラウンド補正あり) の2つの場合について検討した。

$$y - d = a \cdot (x - b)^c \quad (3)$$

Fig.2 の(a1)から(a3)はバックグラウンド補正なし、(b1)から(b3)はバックグラウンド補正ありである (縦軸はそれぞれ、1/2, 1/3, 1/4 になっている)。補正有り・無しともに、シミュレーションで設定したべき乗数 (パラメータ c) の逆数の時 (1/3 乗の時、Fig.2 の(a2), (b2)に該当) べき乗関数部分が直線となり、決定係数が一番良い。なおグラフ内に書かれている、shift は Fitting 結果の b の値、 n はパラメータ c 、であり、 R^2 は全領域での決定係数、 R^2_{bp} は x が shift より大きい領域での決定係数である (バックグラウンドのノイズが大きい場合には、こちらの決定係数を指標として利用できる)。 x が大きい領域 ($x=8$ 以上) で Fitting 線とデータとのずれが生じ、その方向が 1/2 乗 ((a1)と(b1)のグラフ) ではデータが上側に、1/4 乗 ((a3)と(b3)のグラフ) ではデータが下側に来ている。このずれの方向を見れば、べき乗数の程度

を定性的に判断することができる。 ずれの定量的な評価は、Fitting 関数の決定係数を指標として行う。決定係数が最も 1 に近いものは、1/3 乗 ((a2)と(b2)のグラフ) を適用したものである。このことから決定係数を比較することでどのべき乗数の当てはまりが一番良いか評価ができる。傾向として、 n が小さいと shift は大きくなる。バックグラウンド補正なしは補正ありに比べ、shift は大きな値になっている。これは、バックグラウンドも含めてべき乗を適用しているためである。(b2)のグラフ (補正あり) では、shift はシミュレーションで設定した値と一致している。バックグラウンドの値を変えた時の Fitting によって得られる shift 変化を Fig.3 に示す。(なおバックグラウンド以外のパラメータ (a, b, c) は同じである。) Fig.3 において、(a1)と(b1)はべき乗の逆数を適用する前のスペクトルであり、(a2)と(b2) は適用後のスペクトルである。(a1)と(a2) はバックグラウンドが 0 の場合であり、Fitting によって得られた shift は設定した値と一致する。(b1)と(b2)はバックグラウンドをわざわざ 1 にした場合 (バックグラウンドと強度の最大値の強度比を求めると、限りなく 0 である) でも、shift は大幅に高い方に移動している。 Fig.3(c)は(a2)と(b2)などの Fitting 結果を shift と強度比でプロットしたものである。パラメータの d 値は 0, 1, 5, 10, 15, 20, 30, 40, 50 に変えて計算している。縦軸は shift の値であり、横軸はバック

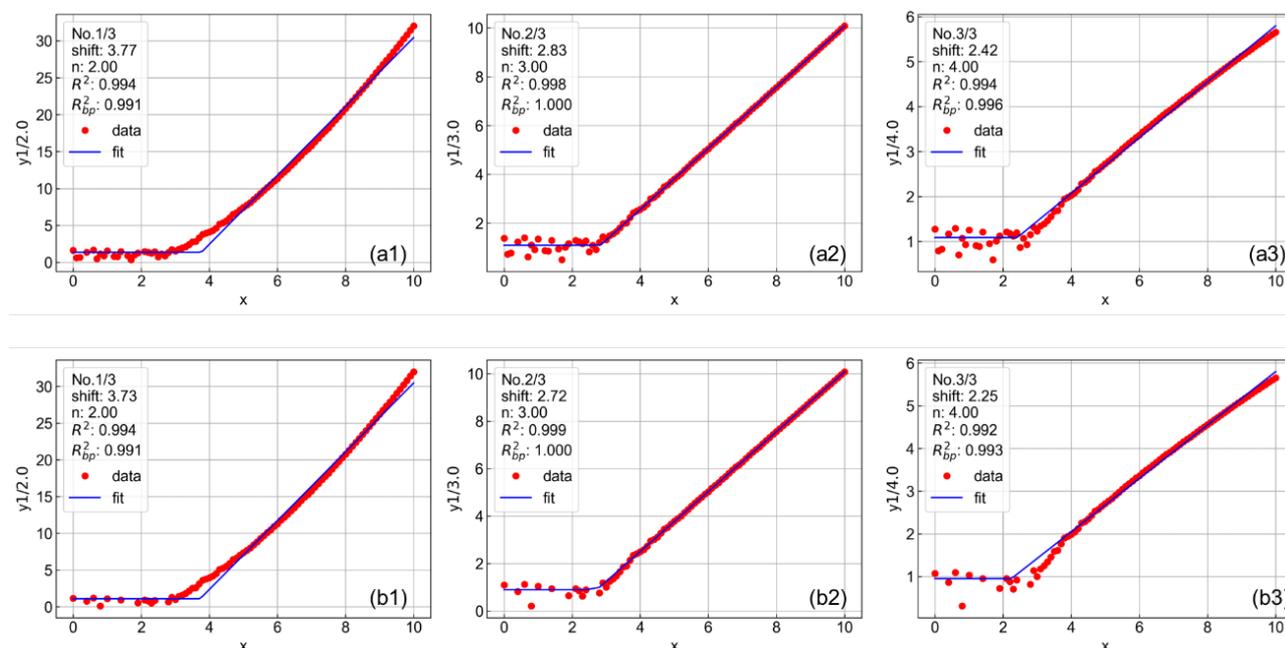


Fig. 4. $1/n$ -Power Plots and their fitting results for spectra with Gaussian noise (mean (μ) = 0, variance (σ^2) = 1) added to the calculated spectra in Fig. 2. (a1), (a2) and (a3) without background correction; (b1), (b2) and (b3) with background correction. Background correction is performed by fitting without background correction and using the background parameters. (color online)

クグラウンドと最大強度と比である。グラフが2本書かれているがこの違いは、一つは、 $1/3$ 乗を適用する前の強度比 (Fig.3 のグラフの(a1), (b1)での強度比)であり、もう一つは、 $1/3$ 乗を適用した強度比 (Fig.3 のグラフの(a2), (b2)での強度比)である。強度比を取ると $1/3$ 乗を適用後の方が適用前よりも比は大きくなる。これは、べき乗を取るとバックグラウンド成分が強調されるためである。比が大きくなると shift も大きい方向へシフトしていくことがわかる。この結果からバックグラウンド補正が行われていない場合、shift は大きく見積もられる。

計算したシミュレーションスペクトルにガウスノイズ (平均値(μ)=0, 分散(σ^2)=1) を付加したスペクトルについても Fig.2 と同様の解析を行った結果を Fig.4 に示す。Fig.4 の(a1)から(a3)は Fig.2 と同様にバックグラウンド補正なし、(b1)から(b3)はバックグラウンド補正ありである。なお、バックグラウンド補正は、最初にバックグラウンド補正なしで Fitting を行いその時のバックグラウンドのパラメータをバックグラウンドの補正值として利用している。補正有り無しともに、Fig.2 と同様に $1/3$ 乗を適用したときに当てはまりが一番良く、shift はノイズがないときに設定した値より大きい値となった。Fig.2 と異なり補正を行っても、shift が設定した値になることはなかった。バックグラウンドの補正がうまくいかずに若干バックグラウンド成分が残っているた

めと考えられる。(なお、シミュレーションにおいてガウス分布を仮定したノイズを加えた場合、バックグラウンドで負の値が生じる場合がある。実際の計測では、マイナスになることはなくその場合0になるが、計算では負の値はそのまま取り扱っている。そのため、べき乗の演算を行うと、負の値は非数となりデータポイントから削除されている。結果として正の値のみが残るためにバックグラウンドレベルが0にはならない。) わずかなバックグラウンドでも Fig.3 に示したように shift は大きい値に移動することから、この効果で設定した値よりも大きくなったと考えることができる。

3. べき乗数の事前情報がない場合の解析方法 (Shift-Log-Log Plot)

3.1 Log-Log Plot (両対数プロット) と Shift-Log-Log Plot

出力値 (計測値) が入力変数に対してべき乗則に従って変化する物理現象のべき乗数の見積もりには、両対数プロットが利用される。一般的な解説として、対象とする関数が

$$y = a \cdot x^c \quad (4)$$

の時、両対数 (自然対数 \ln) を取ると

$$\ln(y) = \ln(a) + c \cdot \ln(x) \quad (5)$$

となる。ここで、 a は定数、 c はべき乗数である。

この式から $Y' = \ln(y)$, $X' = \ln(x)$ とすると、両対

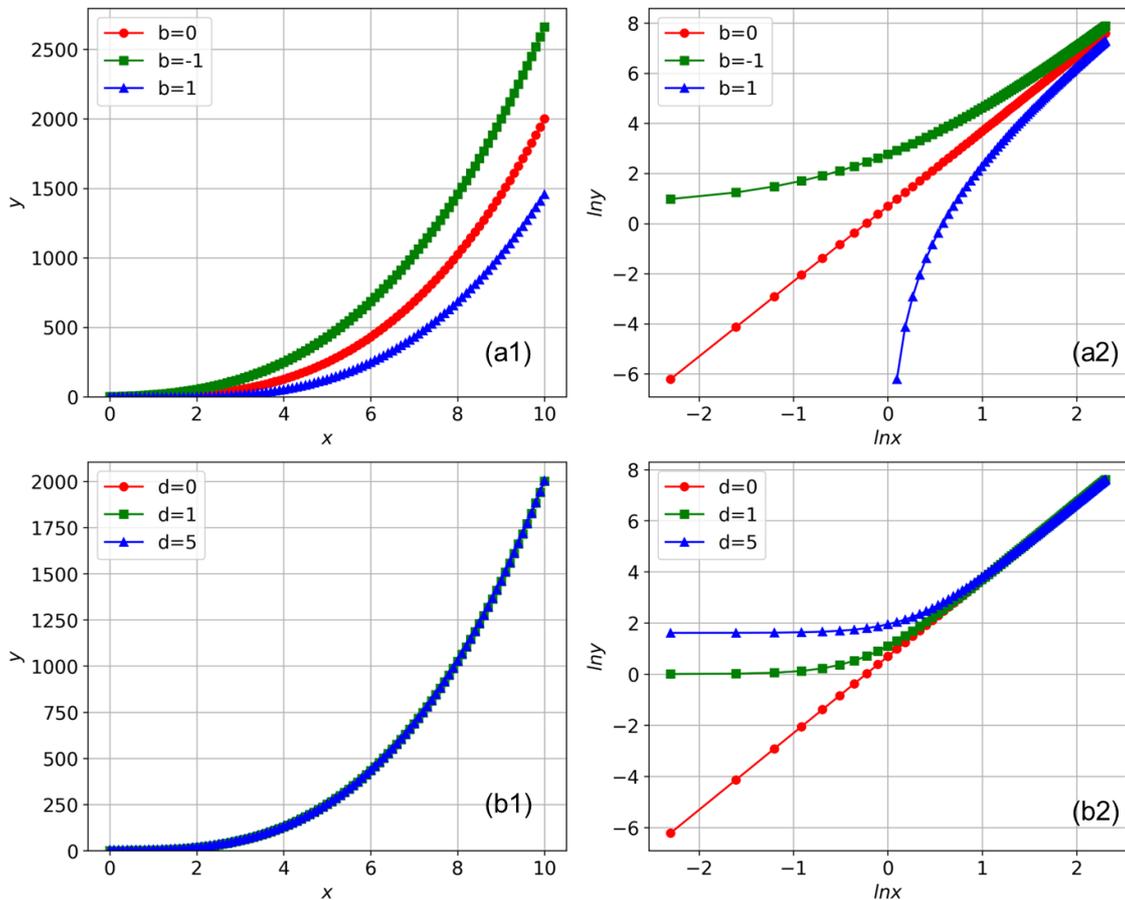


Fig. 5. Spectra calculated for different values of shift and background and their log plot.(a1) is the value of y for different threshold values and (a2) is its log plot. (b1) is the value of y when the background is varied, and (b2) is its log plot. (color online)

数グラフは1次関数とみなすことができ、傾きから、べき乗数を求めることができる。PYSスペクトルのように、出力値がべき乗則に従って増加し、その始まり位置（閾値）が原点から移動している場合、両対数プロットは直線にはならず、べき乗数を正確に求めることができない。(4)で示した式に入力変数方向に移動を取り入れた式は(1)式と同じであり、

(1)式について両対数を取ると

$$\ln(y - d) = \ln(a) + c \cdot \ln(x - b) \quad (6)$$

となる。ここで、 $Y = y - d$ 、 $X = x - b$ とすると上式は、

$$\ln(Y) = \ln(a) + c \cdot \ln(X) \quad (7)$$

となる。 b (shift) と d (バックグラウンド) を変化させたときの式(1)の計算とその両対数プロットをFig.5に示す。Fig.5の(a1)はshift(b)変化の y の値であり、(a2)は両対数を取ったものである。1より小さい値と負の値の対数を取る場合には注意が必要である。1より小さく0より大きい値の対数は0から $-\infty$ の値を取る。負の値は定義されていないことから非数となる。なお、 $-\infty$ も計算では非数の扱いと

なる。グラフにおいても非数となった点は表示されない。shiftが1の時の値(マーカー▲)が、横軸0より小さい値でデータ点がないのは非数となっているためである。shiftが0の時(マーカー●)、両対数プロットは直線となるが、shiftが変わると直線にならない。shiftが負の値の時(マーカー■)には、shiftが0の時の直線より上に位置し、shiftが正の値の時には、下に位置している。両対数のグラフの傾きからべき乗数を求める場合、解析に用いる領域によって傾きが異なる。Fig.5の(b1)はバックグラウンドの値を変化させたときの y の値であり、(b2)はその両対数を取ったものである。(b1)ではバックグラウンドを変えても3つの計算値であり変化が見られないうが、(b2)のグラフではバックグラウンドの変化が顕著に表れる。およそ $\ln(x)=0.7$ よりも大きい領域では、3つの計算値は直線となり一致する。このようにパラメータ b と d の値によって両対数プロットが直線にならないために正確なべき乗数を求めることは難しい。逆に、この値を推定し補償すると両対数プロットは直線になり、べき乗数を求めることがで

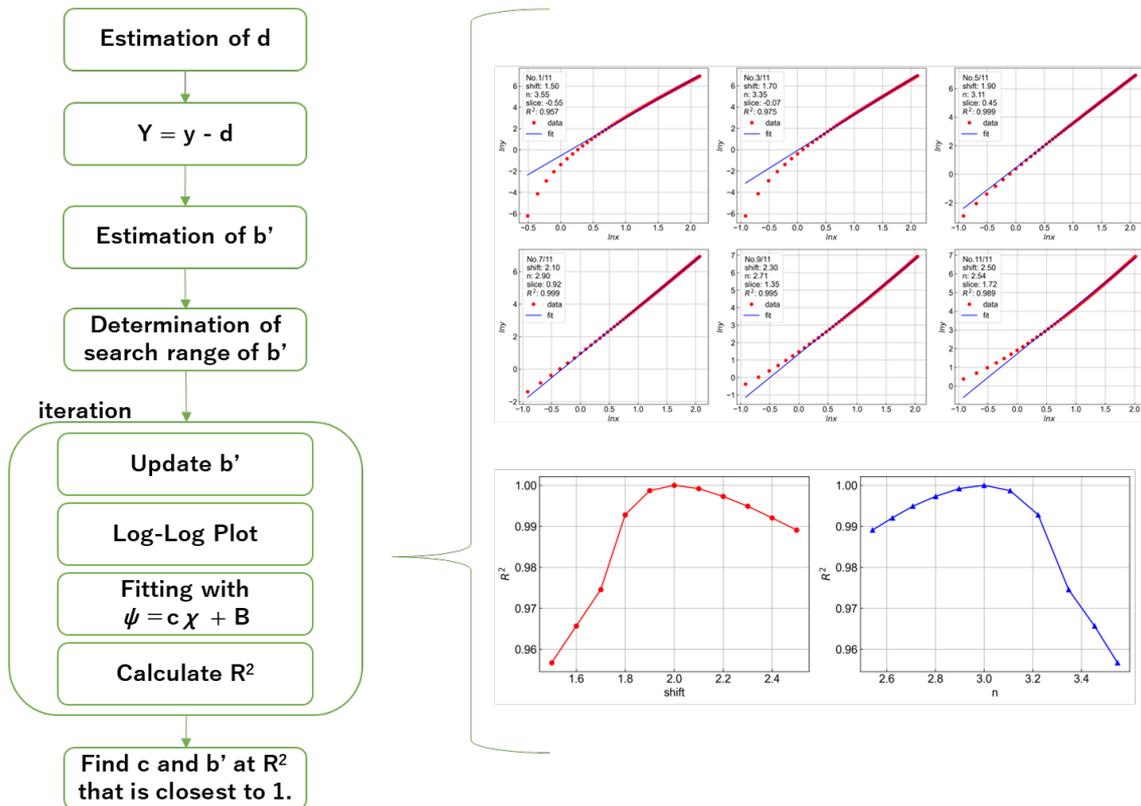


Fig. 6. Analytical algorithm for estimating power and shift (Shift-Log-Log Plot algorithm). (color online)

きる。そこで、最初にバックグラウンド成分を見積もり、その成分を差し引き、 b の値を変えながら（実際には x 軸を移動させる）、両対数プロットを行い、プロットデータが直線になる b を見つけることで、 b と直線の傾きからべき乗数を推定することができる。

3.2 解析アルゴリズム(Shift-Log-Log Plot アルゴリズム)とシミュレーション

前述の通り、適切な b と d で補償後、スペクトルに対して両対数を取ると、スペクトルは1次関数になり c を推定することができる。

ここで、

$$\ln(Y) = \ln(a) + c \cdot \ln(X + b') \quad (8)$$

として b' を変化させたスペクトルに対して、1次関数で Fitting を行いその当てはまり程度を決定係数 (R^2) で評価することによって b 及び、 c を推定する。1次関数における決定係数は相関係数の2乗と等しいことから、決定係数を直線の評価指標として用いる。解析アルゴリズムを Fig.6 に示す。解析アルゴリズムは以下の通りである。

- (1) バックグラウンド(d)を推定する。
- (2) y から d を引く。この値を Y とする。 $Y = y - d$
- (3) 何らかの方法で b' の初期値を推定する。 b' の探

索範囲を決定する。

- (4) b' を変えた $X + b'$ と Y の両対数プロットを作成する。
- (5) 1次関数 $\psi = c\chi + B$ で Fitting を行い、当てはまりを決定係数(R^2)にて評価する。
- (6) 探索範囲内で (4) と (5) を繰り返す。
- (7) 探索終了後、決定係数が1に最も近いときの1次関数の傾き c と、 b' を求める。

3.2.1 ノイズなしの場合のシミュレーション

式 (1) のパラメータ(a, b, c, d)=(2,2,3,1)として計算したスペクトルに対して、3.2 で説明した解析アルゴリズム(Shift-Log-Log Plot アルゴリズム)を適用した結果を Fig.7 に示す。アルゴリズムにおいて、最初にバックグラウンドと shift の初期値及び探索範囲推定を推定する必要がある。バックグラウンドについては、 x の最小値付近における y の値の数点の平均値をバックグラウンドとして、推定することができる。今回のシミュレーションでは、バックグラウンドが1と分かっているのでこの値を用いる。shift の推定方法(後述)はいくつか考えられるが、ここでは初期値がわかっていることとする。探索範囲とステップについては一つの参考基準として、推定初期値を中心に範囲は x の全領域の10%程度、ス

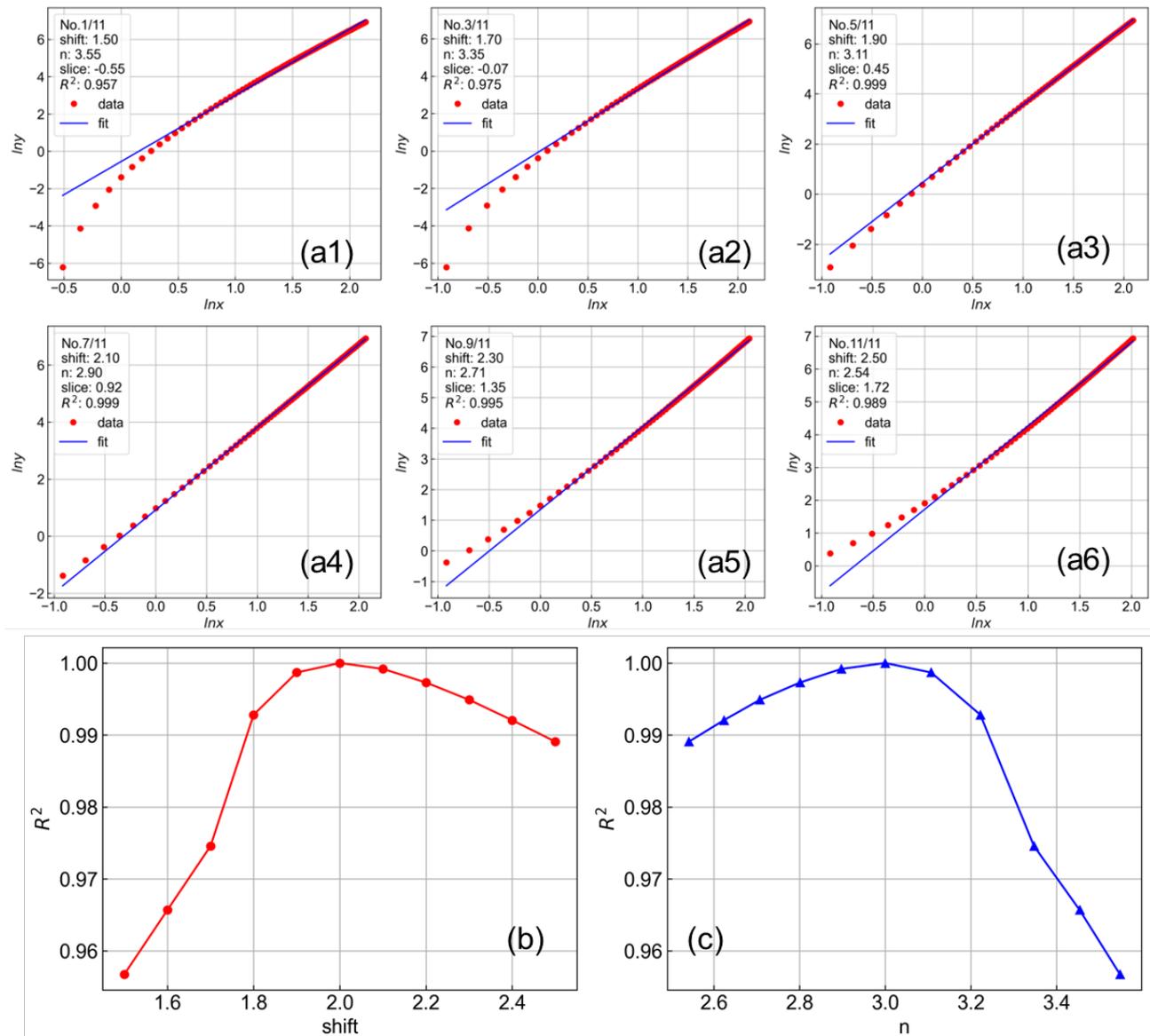


Fig.7: Results of applying the Shift Log-Log Plot algorithm to the computed spectrum. The several graphs from (a1) to (a6) show the Log-Log Plot with varying shift values and the slope evaluated as a linear function. (The search was conducted 11 times, from which the excerpts are shown here). In the legend of the graph, shift (corresponding to parameter b), n (corresponding to parameter c), R^2 (coefficient of determination) for all regions are shown in each figure. (b) and (c) are the results of plotting the shift(threshold) and n (power) obtained in (a1) through (a6) with R^2 . (color online)

テップは、 x のステップと同じ値を用いて探索範囲とする。(なお、ここで x の範囲は0~10なので±0.5、ステップは0.1である。)従って、探索範囲は、推定初期値2を中心に±0.5の範囲で、0.1ステップとする。Fig.7(a1)から(a6)のグラフは、探索閾値(b')を変化させ両対数プロットを作成し、その傾きを1次関数で評価したグラフである。(探索は、11回行っているが、その中から抜粋したもの載せている。左上からNo. 1/11として右へ行くと探索閾値が大きくなる。2段目のグラフも同様である)。なおグラフの凡例には、shift (探索閾値)、 n (Fittingによって得られた傾き)、 R^2 (全領域での決定係数)が記載

されている。探索閾値の増加とともに両対数プロットは徐々に直線になり、その後また直線からずれていく様子がわかる。(a1)のグラフの横軸の最小値が-0.5であり他のグラフの値-1.0と異なっている。これは、探索閾値によってX軸がシフトし、その後、対数を取った影響である。前述の通り、1から0の値と負の値は対数を取ると、それぞれ0から $-\infty$ と非数となる。非数となった点は、グラフ上では表示されないために横軸の値が異なると考えることができる。Fig.7の(b)と(c)は、(a1)から(a6)で得られたshift、べき乗数(n)、決定係数(R^2)を R^2 に対しプロットしたグラフである。グラフから R^2 の最大にお

ける shift は 2 であり, n は 3 であった. Fig.7 の(b) と(c)のグラフにおいて, R^2 の分布が対称形にならない理由は, 前述の通り対数の性質から来ていると考えることができる. 1 より小さい値は対数を取る前の値としては小さいが, 対数を取ると負の大きな値となり, 両対数を取り Fitting を行うと, この値に引きずられる (重みづけが大きくなる) ために当てはまりが悪くなり結果として R^2 の値が小さくなる.

3.3.2 ノイズが乗っている場合のシミュレーション

3.3.1 で用いたスペクトルにガウスノイズ (平均値 (μ)=0, 分散(σ^2)=1) を付加したスペクトルに対して解析アルゴリズムを適用した (Fig.4 の解析で用いた計算シミュレーションと同一のデータ). 探索範囲は, 推定初期値 2 を中心に ± 0.5 の範囲で, 0.1 ステップとし, 3.3.1 と同じ値を用いた. バックグラウンドは, x の最小値付近における y の値の 10 点の平均値とした. なおバックグラウンドの見積もりによって, Fig.5 の(b)に示すように, $\ln(x)$ が 0 より小さい領域で, 1 次関数から大きく外れる原因となり

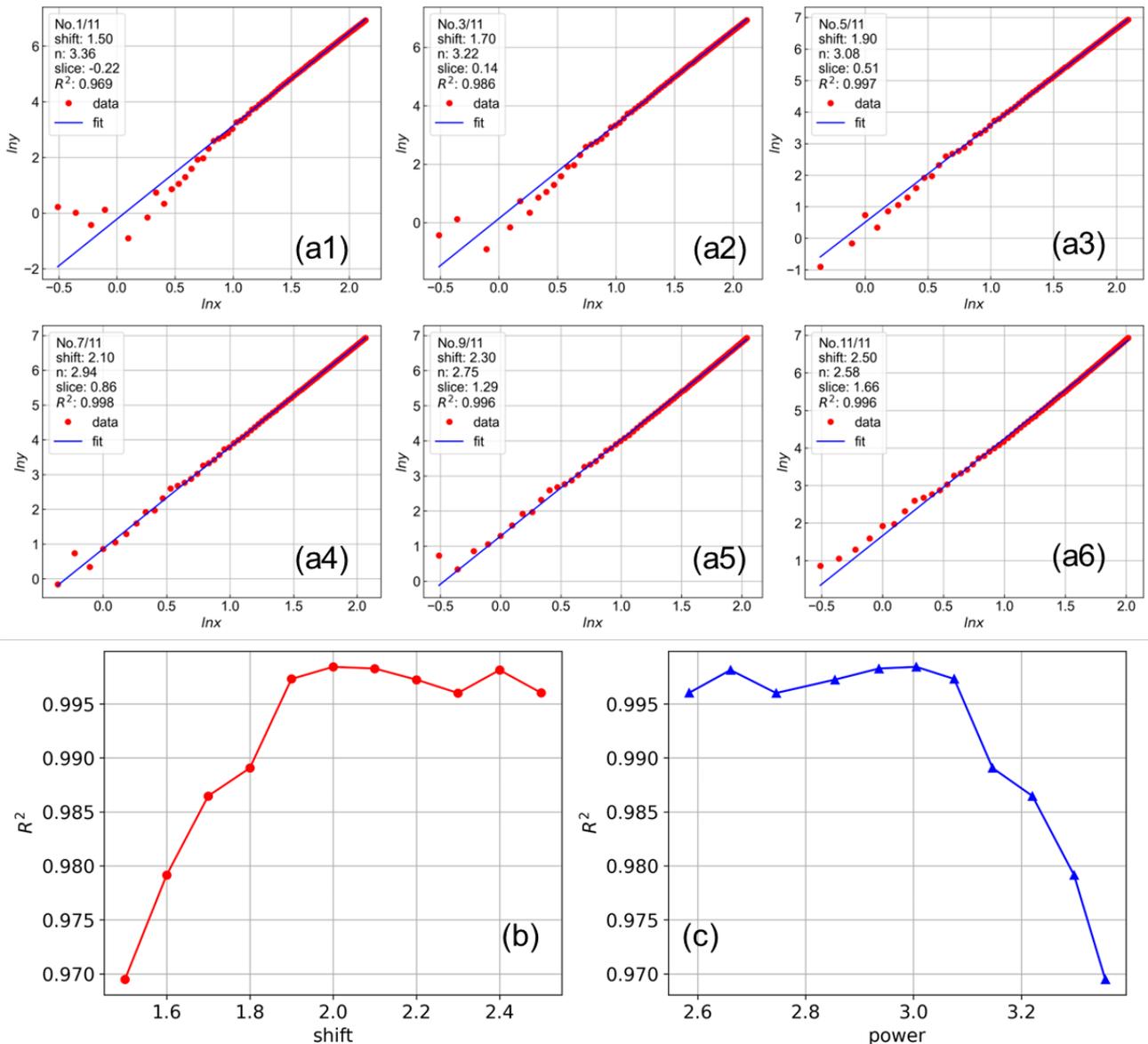


Fig.8: Results of applying Shift-Log-Log Plot to the calculated spectrum in Figure 7 with Gaussian noise (mean (μ) = 0, variance (σ^2) = 1). The multiple graphs in (a) through (a6) are Log-Log Plot with varying shifts, and the slopes are evaluated with a linear function. (The search was conducted 11 times, from which the excerpts are shown here). In the legend of the graph, shift (corresponding to parameter b), n (corresponding to parameter c), R^2 (coefficient of determination for all regions are shown in each figure. (b) and (c) are the results of plotting the shift(threshold) and n (power) obtained in (a1) through (a6) with R^2 . (color online)

傾きの推定に影響が出る. 1 より小さく 0 よりも大きい値に対して対数を取ると実質的な重みが大きくなる. このために 0 から 1 の間にノイズがあるとその影響を大きく受ける. そこで, x 軸の最小の値を決め, それ以上の範囲で Fitting 解析を行うためのパラメータ (margin) を導入した. margin の値は, 自動的に決めることができずノイズなどにもよるが, 大体 $x=0.3$ から 0.5 の範囲を用いている. (なお, $\ln(0.3) \doteq -1.2$, $\ln(0.5) \doteq -0.69$ である.) そして対数の性質上, 負の値の対数は定義されていないために非数となり計算から除外される. margin=0.5 として, Shift-Log-Log 解析結果を Fig.8 に示す. (a1) から (a6) のグラフは Fig.7 と同様の傾向を示している. (b) と (c) の決定係数 (R^2) に対しての shift, べき乗数 (n), をプロットしたグラフにおいて, 決定係数がばたついているのがわかる. これは, x 軸 ($\ln(x)$) の 0 より小さい領域の測定点が影響していると考えられる. R^2 の最大値による推定では, べき乗数は 3.0 となっているが, 同程度の決定係数が複数ありノイズの乗り具合によっては, 推定されるべき乗数が小さくなることも考えられる. ノイズの乗り具合によっては, margin をもう少し大きくすることで改善することもある.

3.3.3 Shift-Log-Log 解析アルゴリズムの利用でのパラメータ推定

Shift-Log-Log 解析方法は, バックグラウンドの推定値や閾値周辺のノイズの影響を大きく受ける. この解析の適用実行には, スペクトルの適用可能性の確認と 4 つの初期値が必要である. 適用できるスペクトルかどうかの確認として, (1) x の増加によって y が右肩上がりにべき乗で増加するスペクトルとなっている. 例えば, 半導体デバイス評価で用いられる電流-電圧スペクトル (IV スペクトル) では, べき乗ではなく指数関数で増加する現象 (ショットキーなど) もある. この場合, 両対数でなく片対数で直線となることからこの解析法では解析できない.

(2) スペクトルの縦軸が 0 から最大数千の範囲, 横軸が 0 から 10 程度で, 正の値を取っている. 例えば, IV スペクトルでは横軸・縦軸の取り方で共に負の値を取る場合がある. また, 自作の測定器などでは, ゼロ点調整がうまくいかず, バックグラウンドの出力値として, 負の値が出力されることもある (AD 変換ボードの出力値など). このような場合, 補正値を加えるか, 絶対値を取る処理が必要である. また, より広い横軸範囲を対象とすることも可能であるが, これに伴い margin の値も大き

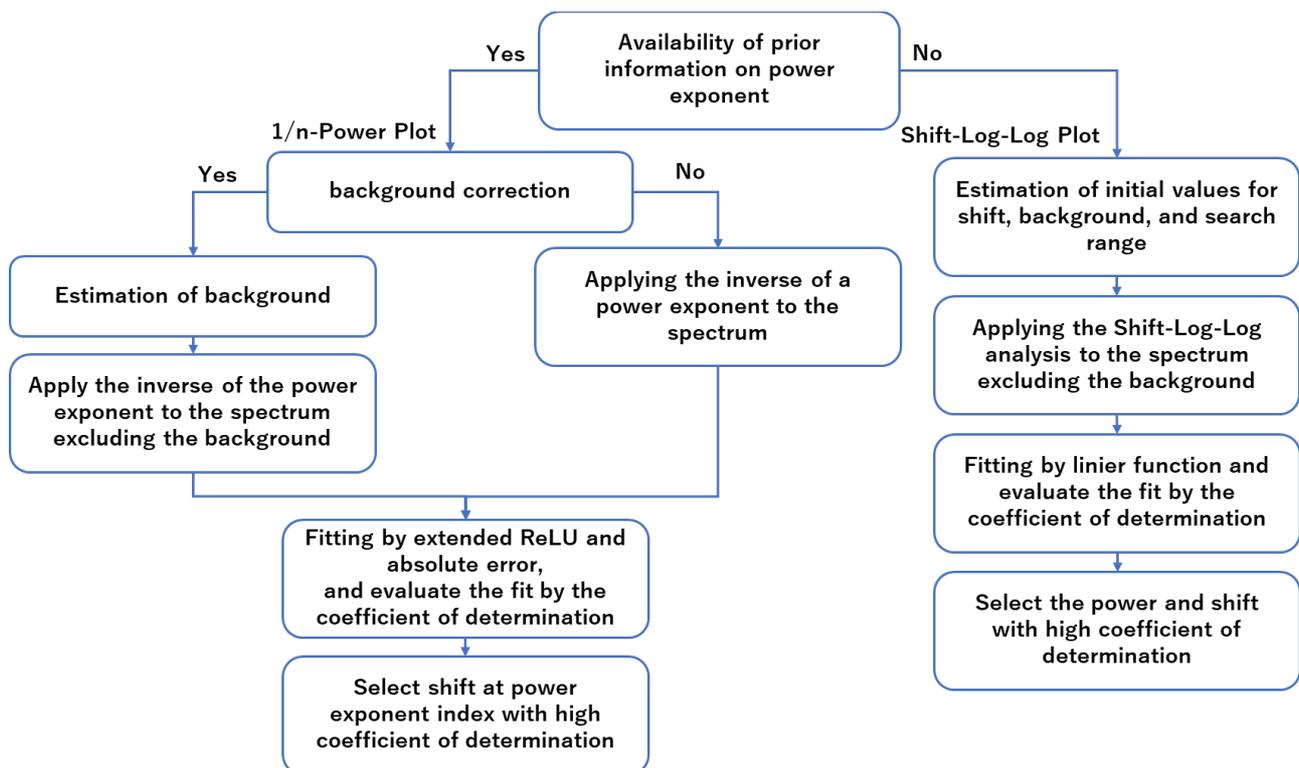


Fig.9: Flowchart of the analysis in the spectrum interpreted by the power law. Depending on the presence or absence of prior information on the power law, the analysis is divided into the 1/n-Power Plot analysis method and the Shift-Log-Log Plot analysis method.

くするなどの推定も必要である。なお、ここで示した横軸の範囲は、我々が本論文で利用したシミュレーション範囲であり、それに伴い margin の初期値も 0.3 から 0.5 としている。初期値として、バックグラウンド, shift, shift を含む探索範囲, margin が必要である。バックグラウンドは、定数のバックグラウンドを仮定し、 x の最小値から近い領域の y の数点を選んで平均化したものを推定値とする。shift と探索範囲の推定については、shift の初期値が指定できない場合は、 x の中間を初期値とし、探索範囲を広くとれば、探索時間はかかるが探索することができる。または、スペクトルに対して $n=2,3$ などの $1/n$ -Power Plot と自動 Fitting を行い、立ち上がり領域を見積もりその値を初期値とする方法もある。探索ステップは、 x のデータのステップ幅を基準とする。margin については、前述の通り我々がシミュレーションで用いた横軸の範囲であれば 0.3~0.5 を基準とし、ノイズの比率（閾値領域のノイズの程度）が多い場合はこの値を大きくする。

4. べき乗解析におけるフローチャート

べき乗則で解釈するスペクトルでは、閾値とべき乗数が重要なパラメータであるとともに、バックグラウンドの見積もりも重要である。理論的なモデルにより予めべき乗数の候補がわかっている時は、 $1/n$ -Power Plot 法を用い、べき乗数についても推定が必要な場合は、Shift-Log-Log Plot 法を用いる。なおこれらの方法は、相補的な関係にある。2 つの方法における解析フローチャートを Fig.9 に示す。事前情報として予めべき乗数の情報があるかどうかで 2 つの方法に分かれる。事前情報がある場合は、 $1/n$ -Power Plot を用いる。さらに、バックグラウンド補正の有無によって分岐する。バックグラウンド補正を行う場合は、何らかの方法でバックグラウンドを推定し、スペクトルからバックグラウンドを差し引く。そしていくつか候補となるべき乗数の逆数を適用したスペクトルを作成し Fitting を行い、決定係数が 1 に近いべき乗数とその時の閾値を求める。バックグラウンド補正を行わない場合は、いくつか候補となるべき乗数の逆数を適用したスペクトルを作成し Fitting を行い、決定係数が 1 に近いべき乗数とその時の閾値を求める。なお、バックグラウンドの推定には、最初にバックグラウンド補正を行わない Fitting を行い、その Fitting パラメータにあるバックグラウンドを用いて、バックグラウンド補正を行うことも可能である。事前にべき乗数がわからな

い場合は、Shift-Log-Log Plot を用いる。初期値（バックグラウンド, shift, shift 探索範囲, margin) の初期値を何らかの方法で推定し、スペクトルからバックグラウンドを差し引き、Shift-Log-Log アルゴリズムを用いて決定係数が 1 に近い shift とべき乗数を求める。初期値の推定に、 $1/n$ -Power Plot を利用して大まかな shift の推定及びバックグラウンドの推定を行うこともできる。例えば、 $1/2$ 乗、 $1/3$ 乗、 $1/4$ 乗で解析を行い、その shift から探索範囲を決定する。べき乗数が求まったところで、 $1/n$ -Power Plot を用いて再度、shift を求めることも可能である。

5. 適用例

5.1 PYS データ

Fig.1 に示した測定スペクトルに対して、 $1/n$ -Power Plot を行った結果を Fig.10 に示す。PYS ではべき乗数の逆数の候補として $1/2$ 、 $1/3$ 乗とする。Fig.10 の(a1)は Au のスペクトルに $1/2$ 乗を適用したもので、(a2)は Au のスペクトルに $1/3$ 乗を適用したものである。(b1)及び(b2)は PTCBI のスペクトルに同様の処理を行ったものである。Fig.1 の手動解析では、バックグラウンド補正は行われていない。Au においては、(a1)グラフでの $1/2$ 乗の決定係数が高く、閾値は手動で解析した値とほぼ一致した(手動解析 5.03 eV に対し自動解析 5.06 eV)。一方で、有機半導体の PTCBI では(b2)グラフの $1/3$ 乗の決定係数が高く、その際の閾値は手動で解析した値とほぼ一致している(手動解析 5.96 eV に対し自動解析 5.98eV)。このことから、有機半導体の解析に際して解析者は以下のように判断したと考えられる。 $1/2$ 乗にしたところ、直線からずれるので、閾値に近いところで線を引いた。これは、 $1/2$ 乗で直線ならずエネルギーの高い領域で強度が増加するので、適用するべき乗数を大きくしなければならないことを意味している。 $1/2$ 乗でも $1/3$ 乗でも近似的に閾値付近では直線性が成立するので、その領域を取ると、実際に $1/3$ 乗した値と同じ結果となった。今回、自動解析方法とその決定係数を指標として比較検討することで、妥当なべき乗数とその際の閾値を自動で数理的に求めることができる。Fitting によって得られたバックグラウンドの値を用いて、バックグラウンド補正の結果は、図には示していないが Au では 4.95eV 有機半導体では 5.93 eV であった。補正によって 0.05 から 0.1 程度小さくなると考えられる。

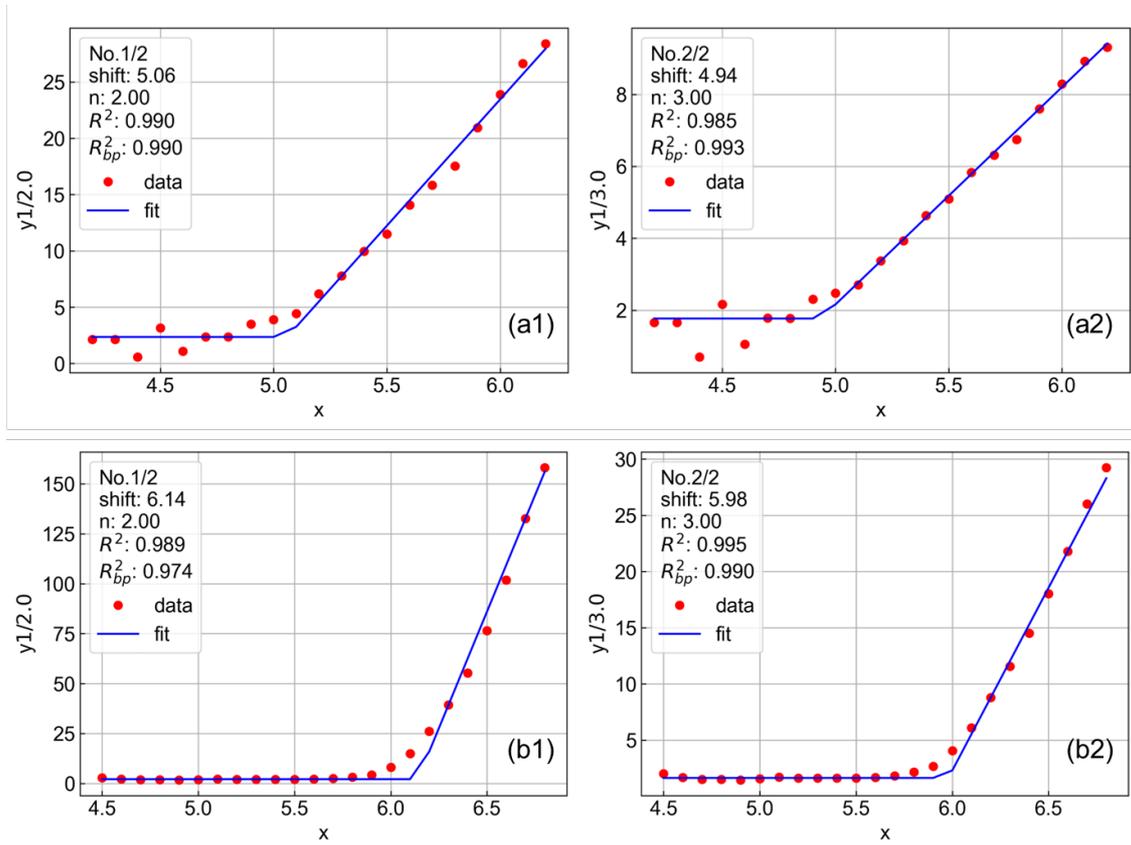


Fig.10: 1/n-Power Plot results for the measured spectrum in Figure 1. The 1/2 power and 1/3 power are used as candidates for the inverse of the multiplier. No background correction was performed in the manual analysis of Figure 1. (a1) is the Au spectrum with the 1/2 power applied and (a2) is the Au spectrum with the 1/3 power applied. (b1) and (b2) are the results of the same process applied to the PTCBI spectrum. For Au, the coefficient of determination of the 1/2 power in graph (a1) is high and the threshold value is almost identical to the manually analysed value (5.03 eV for manual analysis versus 5.06 eV for automatic analysis). On the other hand, for the PTCBI of organic semiconductors, the coefficient of determination of the 1/3 power in graph (b2) is high and the threshold value in this case is almost identical to the value obtained by manual analysis (5.96 eV for manual analysis versus 5.98 eV for automatic analysis). (color online)

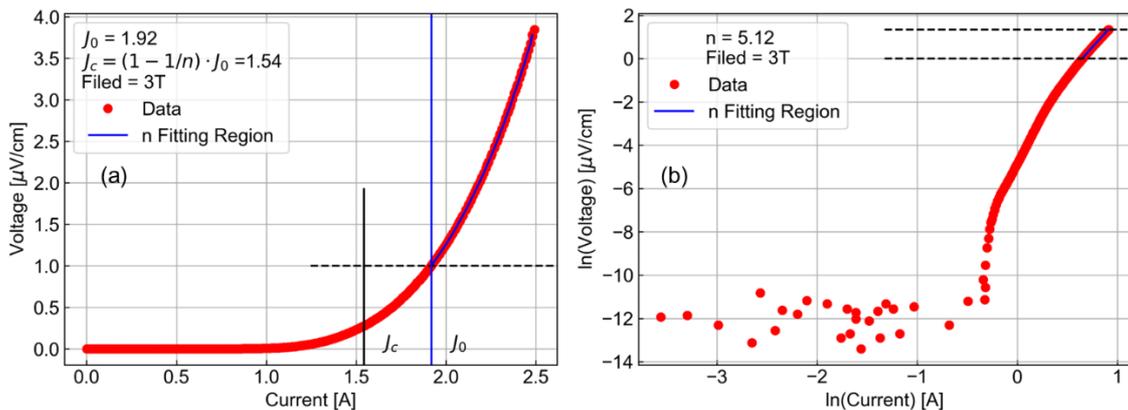


Fig.11: Analysis of critical current density measurement (VI measurement) spectra of superconducting materials using the conventional method. (a) is a linear plot and (b) is a double log-log plot. The critical current density (J_0) is the value of the current density when an electric field of 1 $\mu\text{V}/\text{cm}$ is generated (electric field reference method). n value was calculated from the slope of the log plot of current voltage in the range 1-10 $\mu\text{V}/\text{cm}$. $J_0 = 1.92$, $n = 5.12$. J_c is 1.54, obtained from equation (9). (color online)

5.2 超伝導材料データ

超伝導材料の臨界電流密度測定 (VI 測定) スペクトルでは, ある電圧が生じるまで流せる電流の限界

値が臨界電流密度 (閾値) となり, 非線形性の強さを示す指標が n 値 (べき乗数) である. 測定に用いた超伝導材料は MgB_2 であり温度 4.2 K にて VI 測

定を行った。測定結果および、従来法で解析した結果を Fig.11 に示す。(a)は直線プロット、(b)は両対数プロットである。超伝導の研究解析の歴史は古く、業界標準として、臨界電流密度 (J_0) は $1 \mu\text{V}/\text{cm}$ の電界が発生するときの電流密度の値(電界基準法)、 n 値は $1\sim 10\mu\text{V}/\text{cm}$ 範囲での電流電圧の両対数プロットの傾きから算出することになっている[7]。また、 J_0 とオフセット法(我々が閾値と呼んでいるもの)における臨界電流密度 J_c とは以下の関係がある。

$$J_0 = \frac{J_c}{(1 - 1/n)} \quad (9)$$

測定によって結果導き出された値は、 $J_0 = 1.92$ ($J_c = 1.54$), $n = 5.12$ であった。グラフには解析領域(破線で示された領域)を示しているが、解析領域は固

定されているので値の算出は容易であるが、スペクトル形状を反映した解析を行っているとは言えない。Shift-Log-Log Plot による解析結果を Fig.12 に示す。最初に、初期値を推定するために $n = 2, 3, 4$ として、 $1/n$ -Power Plot を行い、大まかな shift およびバックグラウンドを導き出し、Shift-Log-Log Plot による探索を行った。その結果、shift は 0.78、べき乗数(n) は 3.2 となった。なおこの方法で求めた閾値は J_c であり、式 (9) を用いると、 $J_0 = 1.13$ となった[7]。今回用いた方法と業界標準によって求めた値との直接的な比較ができないが、閾値およびべき乗数は小さく見積もられている。

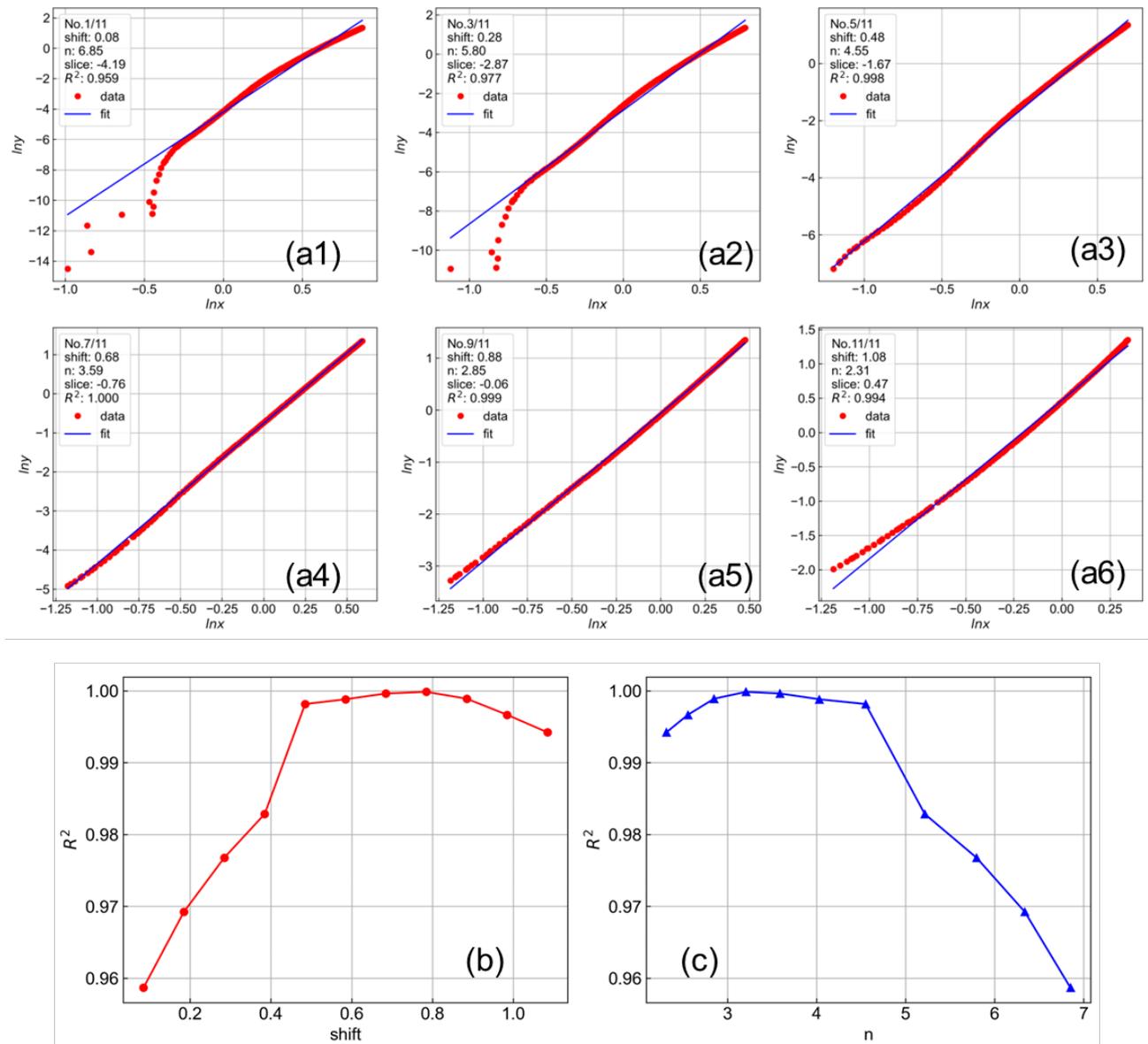


Fig.12. Shift-Log-Log Plot analysis results. The graphs (a1) through (a6) are the results of creating Log-Log Plots with different shift values and evaluating their slopes as linear functions. (b) and (c) are the results of plotting the obtained shift (threshold) and n (power) with the coefficient of determination (R^2). The shift value (J_c) is 0.78 ($J_0 = 1.13$) and the power (n) is 3.2. (color online)

6. まとめ

べき乗則に従って解釈されるスペクトルにおいて、出力値がある入力値を境にべき乗則に従って増加が始まる閾値（原点からの移動量）とそのべき乗数の自動推定方法について、理論的なモデルによる事前情報（べき乗数の値）の有無により2つの方法（1/n-Power Plot 及び、Shift-Log-Log Plot）を提案した。事前情報がある例として、PYS スペクトルの解析例を示し、ない例として、VI スペクトルの解析例を示した。

• 1/n-Power Plot

理論的なモデルによるべき乗数の事前情報が与えられている場合には、このべき乗数の逆数をスペクトルに適用することで、直線が得られる。この直線とバックグラウンドの交点が閾値となる。この交点を自動で求めるために、Fitting 関数として ReLU 関数を拡張した関数を用い、最小化方法として絶対誤差法を用いて最適化 (Fitting) を行うことを提案した。複数のべき乗数の候補がある場合には、それぞれに対し Fitting を行い、決定係数を比較することで、最適なべき乗数とその時の閾値を求めることができる。また、バックグラウンドの補正（バックグラウンドを引いてゼロにする）が行われていない場合、閾値は大きくなる。

• Shift-Log-Log Plot

べき乗数の事前情報がない場合では、両対数プロット (Log-Log Plot) による解析方法に閾値の効果を取り入れた Shift-Log-Log Plot を提案し、そのアルゴリズムについて説明した。閾値による入力値 (x 軸) 方法への移動が起こると Log-Log Plot は直線にはならないが、閾値を補償することで直線となる。直線となるこの閾値を見つけ出しその際の直線の傾きからべき乗数を推定することができる。この直線性の評価は、決定係数にて行う。

• 実測スペクトルへの解析法の適用

実測の PYS スペクトル、VI スペクトルに、それぞれの推定法を適用し、閾値およびべき乗数を自動で求めた。PYS スペクトルでは、金属 (Au) と有機半導体 (PTCBI) 材料について 1/n-Power Plot 解析を行った。PYS ではべき乗数の候補が複数あるためにそれぞれについて Fitting を行いその当てはまり具合について評価した。Au では、1/2 乗の当てはまりが良く PTCBI では 1/3 乗の当てはまりが良かった。当

てはまりが良い乗数における閾値は、手動で求めた閾値とほぼ一致した。VI スペクトルでは Shift-Log-Log Plot にて解析を行い、べき乗数及び閾値を求めた。業界標準の従来法では、スペクトル形状を無視した解析法であり、本手法で求めた値と比較することができなかった。従来法に比べ、閾値、べき乗数ともに小さい値となった。

7. Reference

- [1] R. Shimizu, S. Kobayashi, Y. Watanabe, Y. Ando, T. Hitosugi, *APL Mater* **8**, 111110 (2020). (doi: 10.1063/5.0020370)
- [2] 篠塚寛志, *表面と真空* **65** (3), 121 (2022). (doi: 10.1380/vss.65.121)
- [3] 柳生進二郎, 吉武道子, 知京豊裕, 長田貴弘, *表面と真空* **63** (6), 270 (2020) . (doi: 10.1380/vss.63.270)
- [4] 柳生進二郎, 松本明善, 松波成行, 西島元, 伴野信哉, 石井秋光, 吉武道子, 長田貴弘, ”べき乗則に従うスペクトルの解析” 第 70 回応用物理学会春季学術講演会 17a-PB02-4 (2023).
- [5] J. Tauc, R. Grigorovici, and A. Vancu, *Phys. Status Solidi*. **15** 627 (1966).
- [6] H. Fowler: *Phys. Rev.* **38**, 45 (1931).
- [7] E.O. Kane: *Phys. Rev.* **127**, 131 (1962).
- [8] M. Kochi, Y. Harada, T. Hirooka and H. Inokuchi: *Bulletin of the Chemical Society of Japan* **43**, 2690 (1970).
- [9] H. Ishii, D. Tsunami, T. Suenaga, N. Sato, Y. Kimura, M. Niwano: *Hyomen Kagaku* **28** [5] 264 (2007). (doi: 10.1380/jsssj.28.264)
- [10] [https://en.wikipedia.org/wiki/Rectifier_\(neural_networks\)](https://en.wikipedia.org/wiki/Rectifier_(neural_networks)) (閲覧 2023/12/4)
- [11] S. Tanaka, *Oyo Buturi* **59** [5] 647 (1990).

8. Appendix

• 解析コードについて

本論文で用いた解析コードおよびグラフは、github にて本論文が掲載可の後に公開する。

査読コメント, 質疑応答

査読者 1 井上 雅彦 (摂南大)

本論文は、ベキ乗則に従うスペクトルのベキ乗数と閾値に対して、解析者によらず同じ値が得られる自動推定法を提案するもので、表面分析の標準化という観点で重要であり、JSA に掲載する価値があることを認めます。論文をわかりやすくするために修正・追記の提案をいたしますのでご検討ください。

[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。

ご指摘を踏まえ改訂いたしました。

[査読者 1-1]

3.3.3 節のところ、Shift-Log-Log 法が適用できるかどうかの確認の(2)番目、スペクトルの縦軸横軸のオーダーの制約がある理由がよくわかりません。また、オーダーが外れる場合、「掛け算などを行い・・・」とありますが、この操作も自動化して解析手法に含めることはできないのでしょうか？もしそれが可能なら解析者は(2)を考慮する必要がなくなるように思いますがいかがでしょうか？この辺りを捕捉解説していただけると良いと思います。

[著者]

ご指摘の(2)の条件について再度検討したところ、値が負の値を取らないこと、横軸範囲については、特に制限を設ける必要がないこと、に修正いたしました。特に横軸の範囲の制限は margin の値の初期値からきており、横軸の範囲を広げた場合、margin の値を大きくすることで対応することができます。また、Shift-Log-Log 法で解析を行う前まで前処理部分については、各測定スペクトルの特徴に由来することなどから、個別に対応することになると考えています。(Shift-Log-Log 解析を行う関数とその関数にデータを渡す前処理関数に分けて対応する。)

査読者 2 眞田 則明 (アルバック・ファイ)

閾値測定は表面分析でも多くの場面で利用されており、ご指摘のある紫外光電子分光法によるイオン化ポテンシャル測定だけでなく、試料吸収電流測定や(分解能が非常に高くできる場合に限定されます

が)逆光電子分光法による電子親和力測定など、多岐に亘って関連します。JSA 誌に掲載し、広く表面分析関係者に問う価値のある論文と考えます。

なお、以下の点について著者のご意見をうかがえますと幸いです。

[査読者 2-1]

バックグラウンドが外部の擾乱であることが明確で、しかも縦軸のゼロが信頼できる場合は、バックグラウンドを考慮した上で、計測スペクトルと縦軸のゼロとの交点を閾値とする場合があるかもしれません。作成されたソフトウェアで、直線バックグラウンドとの交点ではなく縦軸のゼロ点との交点を閾値として計算することを指示することができるでしょうか？

[著者]

今回の論文で作成したコードにはご指摘のような機能は入っておりません。解析関数に、「縦軸のゼロ点の値を求める処理」を入れることは可能です。一方でご指摘の条件を人間が判断するか、それとも自動で判断するかで難易度が変わってきます。特に自動で判断する場合、「どの値を見てどういった指標で評価し判定するか」について検討が必要だと思います。現状、Fig.9 のアルゴリズムでも示す通り、バックグラウンド補正処理が、ご指摘の処理の値に近いと考えられます。

また、PYS 装置を発売している理研計器では、解析領域(バックグラウンドと直線増加領域)を手動で指定する必要がありますが、バックグラウンドを差し引いて、縦軸のゼロ点との交点を求める処理を、装置に付随する解析ソフトで提供し、その処理結果が出力データに含まれています。なお、その処理によって推定された閾値やデータの可視化は以前 JSA にて発表した文献にコードが記載されています。

柳生 進二郎. 大気中光電子収量分光装置(理研計器 AC シリーズ)の計測ファイル用メタデータ抽出プログラム. *J. Surf. Anal.* **29**, 97 (2022). (doi:10.1384/jsa.29.97).