J. Korean Inst. Electr. Electron. Mater. Eng.
Vol. 35, No. 3, pp. 223-231 May 2022
DOI: https://doi.org/10.4313/JKEM.2022.35.3.3
ISSN 1226-7945(Print), 2288-3258(Online)

Tutorial Status Report

X선 기반 분광광도계를 통해 얻은 데이터 분석의 기초

조재현, 조 욱

울산과학기술원 신소재공학과 및 JULIA 연구센터

초록: 분광학은 재료의 결정학적, 화학적 구조를 분석하기 위해 가장 보편적으로 활용되는 분석 기법이다. 이러한 기조에 따라 다양한 분석 소프트웨어와 peak fitting과 관련된 기술적 가이드라인이 보급되었지만, 정작 '왜' 중간 계산 과정을 거치고 해당 함수를 쓰는지에 대한 논의는 부족한 실정이다. 따라서 본 tutorial에서는 X선 기반 분광광도계를 통해 얻은 데이터 분석의 기초를 논하고자 한다. 이를 위해 관련된 peak fitting을 위해 필요한 실용적 배경지식을 제시하였다. 나아가, 하나의 예시로 임의로 선정한 X선 광전자 분광법 데이터에 대한 curve fitting 과정을 순서에 따라 알기 쉽게 소개하였다. 제시한 기초 이론은 특정 소프트웨어에 국한된 내용이 아니라 fitting tool이 있는 모든 소프트웨어에서 그대로 활용 가능할뿐더러 다른 분광법 데이터를 분석하는 데 적용 가능하기 때문에, 본 내용을 숙지한다면 보다 수월한 연구 진행을 위한 바탕이 될 수 있을 것이라 기대한다.

키워드: X선 광전자 분광학, 가우스 함수, 로렌츠 함수, Voigt 함수

Practical Guide to X-ray Spectroscopic Data Analysis

Jae-Hyeon Cho and Wook Jo

Department of Materials Science and Engineering & Jülich-UNIST Joint Leading Institute for Advanced Energy Research (JULIA), Ulsan National Institute of Science and Technology (UNIST), Ulsan 44919, Korea

(Received March 13, 2022; Accepted March 15, 2022)

Abstract: Spectroscopies are the most widely used for understanding the crystallographic, chemical, and physical aspects of materials; therefore, numerous commercial and non-commercial software have been introduced to help researchers better handling their spectroscopic data. However, not many researchers, especially early-stage ones, have a proper background knowledge on the choice of fitting functions and a technique for actual fitting, although the essence of such data analysis is peak fitting. In this regard, we present a practical guide for peak fitting for data analysis. We start with a basic-level theoretical background why and how a certain protocol for peak fitting works, followed by a step-by-step visualized demonstration how an actual fitting is performed. We expect that this contribution is sure to help many active researchers in the discipline of materials science better handle their spectroscopic data.

Keywords: X-ray photoelectron spectroscopy, Gaussian function, Lorentzian function, Voigt function

✉ **Wook Jo; wookjo@unist.ac.kr**

Copyright ©2022 KIEEME. All rights reserved.
This is an Open-Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (http://creativecommons.org/licenses/by-nc/3.0) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

1. 서 론

■ 분광학(spectroscopy)은 파장에 따른 전자기파와 물질의 상호작용을 연구하는 학문으로, 그 상호작용은 분

광광도계(spectrometers)를 통해 기록되며 재료의 구조적 특성 및 물리적 성질의 원인을 규명할 때 가장 보편적으로 사용하는 분석 기법이다.

- X선 회절(XRD, X-ray diffraction), 라만 분광법(Raman spectroscopy), X선 광전자 분광법(XPS, X-ray photoelectron spectroscopy) 등은 재료 분야 연구에서 수시로 활용되는 분석기법이다.
- X선 회절의 경우 타의 추종을 불허할 정도로 사용 빈도가 높은 까닭에 profile fitting에 특화된 상용 소프트웨어는 물론 무료로 구할 수 있는 다양한 소프트웨어가 있다. 그럼에도 불구하고 많은 초입자들이 실험 데이터에서 원하는 정보를 실제적으로 어떻게 도출하는지에 대해 잘 알지 못하고, 기술적으로는 알고 있더라도 왜 특정 함수를 활용해서 데이터 fitting을 진행하는지에 대해 모르는 경우가 많다.
- 본 논문에서는 분광광도계를 통해 습득한 데이터의 curve fitting에 대한 간단한 기초 이론과 함께 분야 초입자들이 보다 쉽게 데이터 분석에 대한 이해도를 높일 수 있도록 실제 XPS 데이터의 curve fitting이 어떻게 진행되는지에 대한 실례를 제시하고자 한다. 본 논문에서 제시하는 curve fitting은 편의상 국내 재료 분야 연구자들이 가장 많이 사용하는 상용 소프트웨어인 origin Pro™를 사용하였지만, Matlab™, Python™ 등의 소프트웨어가 제공하는 curve fitting 기능을 이용해도 동일한 결과를 얻을 수 있다.

2. 이 론

- 주지하듯이 모든 물질은 원자로 구성되어 있고, 원자는 양성자, 중성자로 구성된 핵과 전자로 이루어져 있다. X선 기반 분석은 이들 원자를 구성하고 있는 전자와 X선 사이의 상호작용을 분석함으로써 물질 내 전자의 고유한 에너지 상태를 해석하는 것이다.
- 절대온도 0도를 제외한 모든 온도에서는 엔트로피 효과로 인해 에너지 준위에도 어느 정도의 불확실성이 존재한다. 즉, 외부에서 조사된 X선을 흡수한 전자가 들뜬 상태로 여기된 후 다시 바닥 상태로 돌아갈 때, 두 에너지 준위 차이에 해당하는 X선을 방출하는데, 이는 해당 에너지 준위 차이가 하나의 정해진 값을 가지는 것이 아니라 분포를 가진다는 것을 의미한다.
- 에너지 값의 분포에 영향을 주는 요인은 다양하지만, 통상적으로 이들 요인의 전체적 결과는 정확한 값을 기준으로 대칭성을 가지기 때문에 평균값 기준 좌우 대칭인 함수를 fitting 함수로 사용한다.
- 일반적으로 실험 데이터에 명백한 비대칭성이 존재하더라도 background noise를 제거하고 중첩된 peak을 분리하면 대칭성이 확보된다.
- 다만, 아직 학계에서도 논란이 많듯이 상기 과정을 통해서도 제거 불가능한 지수적 감쇄에 의한 비대칭성이 간혹 발생하긴 하지만 [1], 최종 결과물에 대한 재료공학적 해석에 있어서는 오차 범위 내에서 큰 영향이 없기 때문에 현 논문에서는 분광광도계로 얻은 데이터의 세부 요소가 항시적으로 대칭성을 가지고 있다는 전제하에 논의를 진행한다.
- 좌우 대칭성을 가지는 peak 함수는 주지하다시피 가우스(Gaussian) 함수와 로렌츠(Lorentzian) 함수로 대표된다. 실제로 두 함수가 fitting에 가장 많이 사용되고 있는데, 많은 경우 두 함수의 기본 특성에 대한 고려가 부족한 탓에 각각의 함수를 부적절하게 사용하여 결과 해석에서 문제가 발생하기도 한다 [2].
- Gaussian, 즉, 정규분포와 Lorentzian의 가장 근본적인 차이는 표준편차의 유무에 있다.
- 다트(darts)를 하는 사람은 다트촉을 중앙에 맞추려 애쓰기 때문에 중앙 부근에서 다트촉의 밀도가 가장 높고 중앙에서 멀어질수록 밀도가 점점 낮아지는 경향이 있는데, 시행 횟수가 충분이 많다면 다트촉의 밀도함수는 Gaussian을 따르게 된다.
- 반면, Lorentzian의 대표적 예는 학부과정 중 확산에 대해 배울 때 가장 먼저 접하는 random walk이다. 술에 취한 사람이 비틀거리며 걷는 경우라면 매번 움직일 때마다 출발점 주위를 맴돌 확률이 높지만 계속된 반복적 움직임으로 출발점에서 무한히 멀어질 수도 있는데 이는 술에 취한 사람이 원래 위치로 돌아가려는 목적의식이 없기 때문이다. 즉, Lorentzian의 경우 시작점이라는 평균값(mean value)은 있지만, 표준 편차는 존재하지 않는 상황을 묘사한다.
- 실재로는 아주 복잡한 이유가 있지만, 간단히 말해서 재료 내 특정 에너지 준위가 엔트로피 효과에 의해 분포를 가지는 경우 앞서 언급한 random walk과 비슷한 상황이므로 결과적인 peak의 모양은 Lorentzian을 따른다 [1,3,4].
- 반면 재료에서 방출되는 X-ray가 검출기(detector)에 포획될 때도 어느 정도의 분포를 가지게 되는데 이는 앞서 예시한 다트를 던지는 상황처럼 Gaussian을 따르게 된다.
- 결과적으로 우리가 분광광도계를 통해 얻는 데이터의 peak은 Gaussian과 Lorentzian 분포가 중첩

(convoluted)된 결과인 셈이다.

- 수학적으로 Gaussian과 Lorentzian 함수를 convolution 하여 모사할 수 있는데, 이 중첩 함수가 Voigt 함수이다. 즉, Voigt 함수를 이용하면 분광광도계를 통해 얻은 데이터의 모든 peak fitting이 가능하다는 의미이다.
- 다만, Voigt 함수는 적분 구간이 음의 무한대에서 양의 무한대까지인 이상적분(improper integral)이 포함되어 있어서 수작업으로 계산하는 것이 거의 불가능에 가까울 뿐만 아니라, 과거 컴퓨터의 성능이 우수하지 못한 시절에는 방대한 계산량으로 인해 실질적으로 활용할 방법이 없었다. 이렇기 때문에, 그 대안으로 정확도가 떨어지는 것은 물론 물리적 근거가 부족한 pseudo-Voigt 함수, Pearson VII 함수 등 다양한 함수가 활용되었다. 하지만 현재의 우수한 컴퓨터의 성능을 고려할 때, 향후 peak fitting에는 고민 없이 Voigt 함수를 사용할 것을 권한다.

3. 본 문

- 전술한 바와 같이 오리진(origin Pro™: origin basic version의 경우 본 tutorial에서 활용한 peak analyzer가 없음을 주지한다) 프로그램은 다양한 분야에서 raw 데이터를 정리하는 데 널리 사용되고 있으며, 데이터 시그널을 분석할 수 있도록 다양한 기능을 제공한다. 따라서 본 tutorial status report에서는 오리진 프로그램을 활용하여, XPS profile 분석 과정의 실례를 step-by-step 방식으로 보여주고자 한다.
- 해당 분석은 본 연구진이 실제로 측정한 ferrite계 세라믹 내 존재하는 Mn 이온에 대한 XPS raw 데이터를 활용하였으며, 2, 3, 4가의 Mn 이온이 공존한다는 가정하에 fitting을 진행하였다.
- 그림 1에 정리된 순서도와 같이, 측정된 데이터의 background noise를 제거하기 위해 baseline을 설정한 후, raw 데이터에서 해당 영역을 제거해 주는 과정

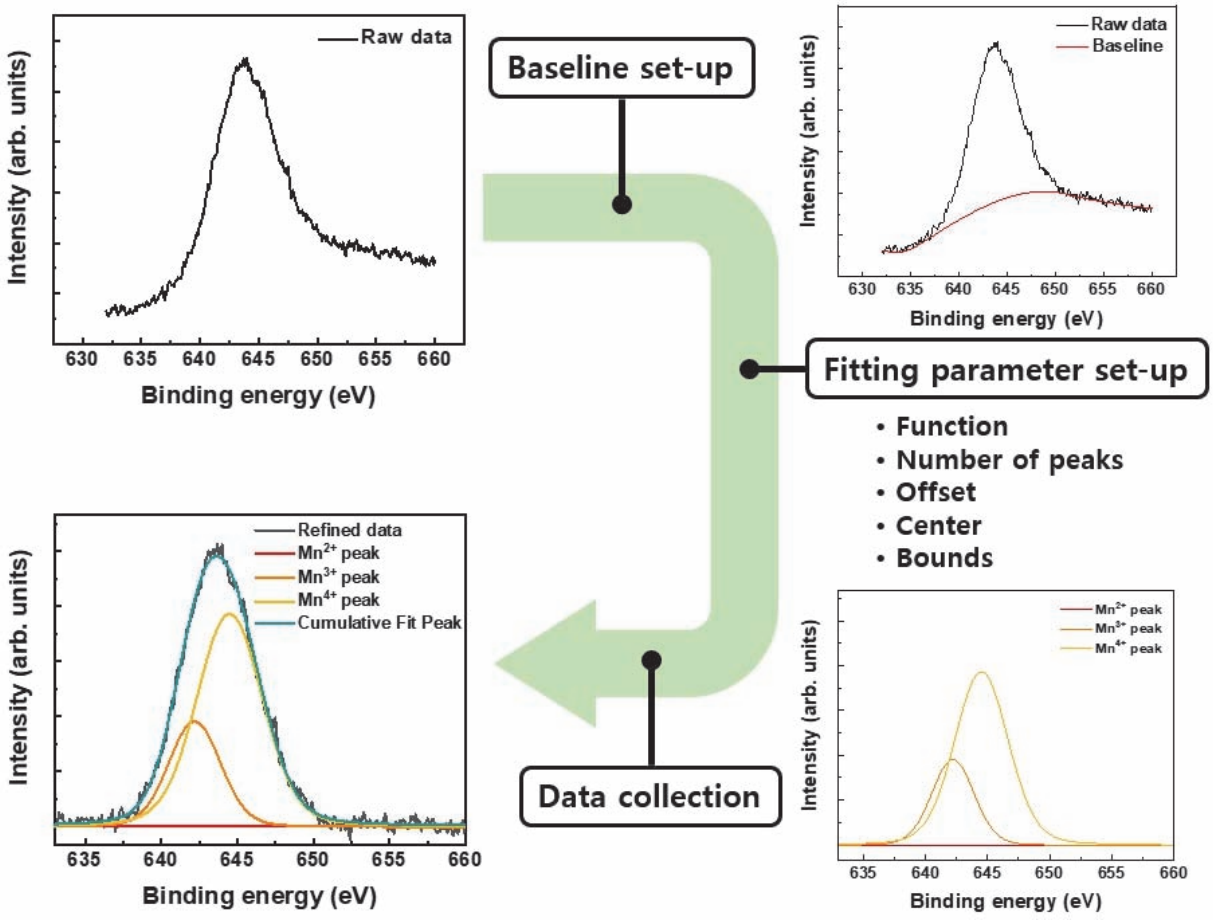

Fig. 1. A flowchart for X-ray photoelectron spectroscopy (XPS) fitting.

을 우선적으로 진행하였다. 그 후, Voigt 함수를 활용하여 Mn 이온의 각 산화수에 해당하는 peak을 구별해내기 위한 fitting을 진행하였다.

- Raw 데이터에서 실제 Mn 이온으로 인한 측정값만을 고려하여 정확한 분석을 하기 위해서는 시편 준비 및 측정 장비 등으로 인한 외부 변수를 제거해야 한다. 이를 위해 측정값에서 baseline을 도출 후 해당 영역을 제거하는 과정을 거쳐야 한다.
- 물론 가장 정확한 방법은 사용하는 장비에서 실측한 background data를 이용하여 background noise를 제거하는 것이지만, 많은 경우 background noise가 최종 결과에 큰 영향을 주지 않기 때문에, 통상 아래에 소개하는 것과 같이 소프트웨어적으로 제거한다.
- Baseline을 설정하기 위해서는 메뉴의 'Analysis'-'Peaks and Baseline'-'Peak Analyzer'-'Open Dialog'로 들어가 'Peak Analyzer' 창을 열어야 한다. 그림 2와 같이, 해당 창에서 'Create Baseline'을 goal로 설정하고 다음 단계로 넘어간다.
- Raw 데이터에 맞는 baseline을 정확하게 설정하기 위해서는 'Baseline Anchor Points'와 'Interpolation Method'를 알맞게 지정해 주어야 한다 (그림 3).
- 'Baseline Anchor Points'는 baseline을 fitting하기 위한 기준점을 잡아 주는 것으로, raw 데이터의 양쪽 밑단 선을 따라 추가 및 위치 수정을 진행한다.
- 'Interpolation Method'는 지정된 'Baseline Anchor Points'를 기반으로 baseline을 계산하는 방식으로, 주로 직선이 아닌 flexible bands의 합으로 나타내는 'Spline'이나 'BSpline (basis spline)'을 사용하는 것이 좋은데 fitting 진행자의 주관에 따라 합리적인 선택이 필요한 부분이다.
- 이러한 과정은 그림 4의 preview 결과값과 같이, 시작점과 끝점이 평탄한 그래프를 얻을 때까지 진행한다. Baseline fitting이 끝난 후, sheet에 도출되어 나온 해당 값을 raw 데이터에서 뺀 뒤 새로운 그래프를 그린다. 해당 과정을 통해 background noise가 제거되었기 때문에, 얻어진 refined 데이터는 신뢰도 높은 peak fitting이 가능하게 된다.
- Background를 제거한 데이터를 이용해 실제 fitting을 진행하기 위해서는 해당 데이터를 선택한 뒤, 메뉴의 'Analysis'-'Fitting'-'Nonlinear Curve Fit'-'Open Dialog'로 들어가 'NLfit' 창을 불러온다.
- 그림 5와 같이, 'Settings'에서 앞으로 진행할 계산의 근간이 될 함수를 선택할 수 있는데, 전술한 바와 같이 여기서 Voigt 함수를 선택하면 된다.
- 함수 선정이 끝나면, 'Advanced' 부분에서 기준이 될 peak의 개수를 지정해 준다.
- 본 report에서는 전술한 바와 같이 Mn의 XPS 데이터를 활용하였고, 산화수(2+, 3+, 4+)에 따른 3가지 개별 peak이 예상되기 때문에 3개의 peak을 fitting하도록 설정하였다.

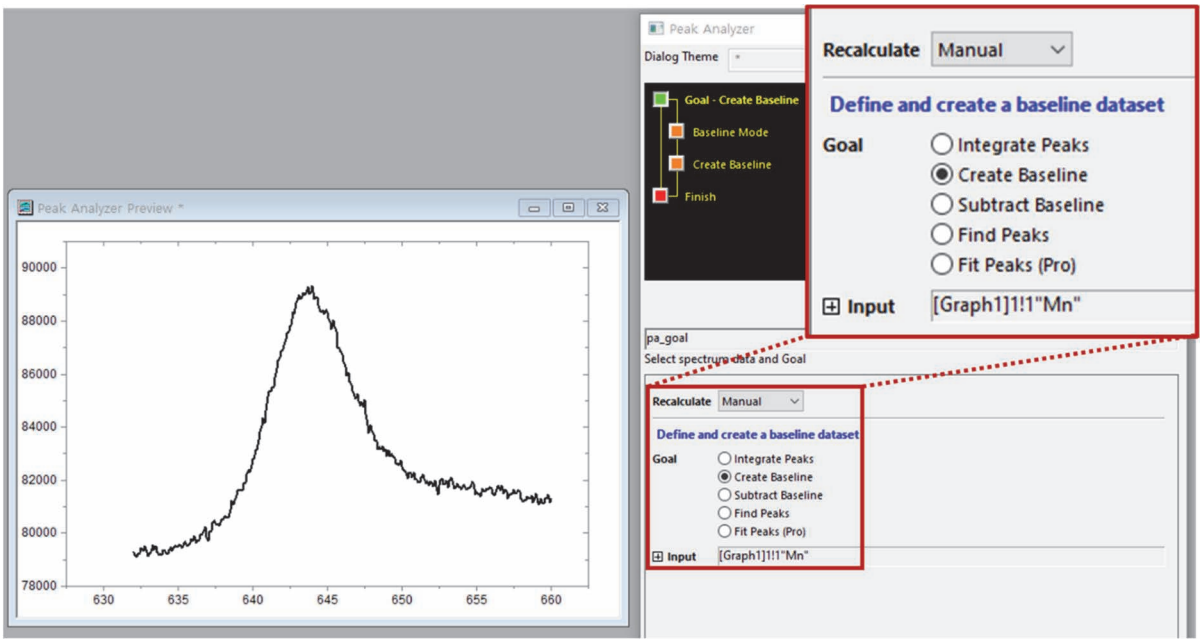

Fig. 2. Goal set-up to create baseline in the peak analyzer tool.

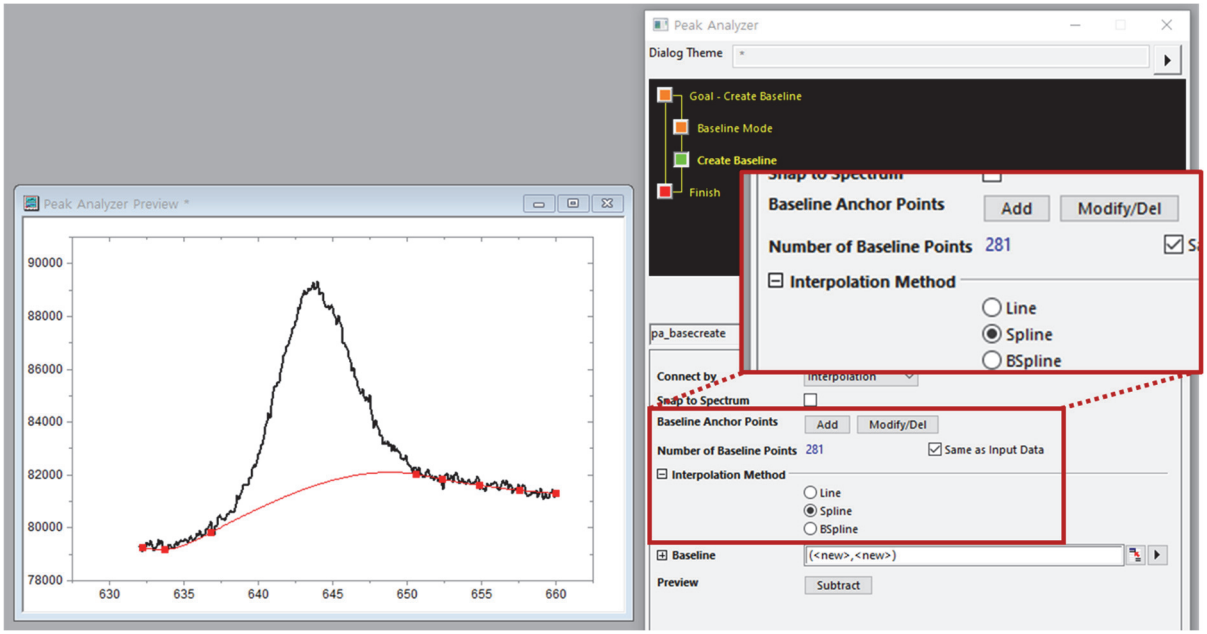

Fig. 3. An example for setting up anchor points and interpolation method.

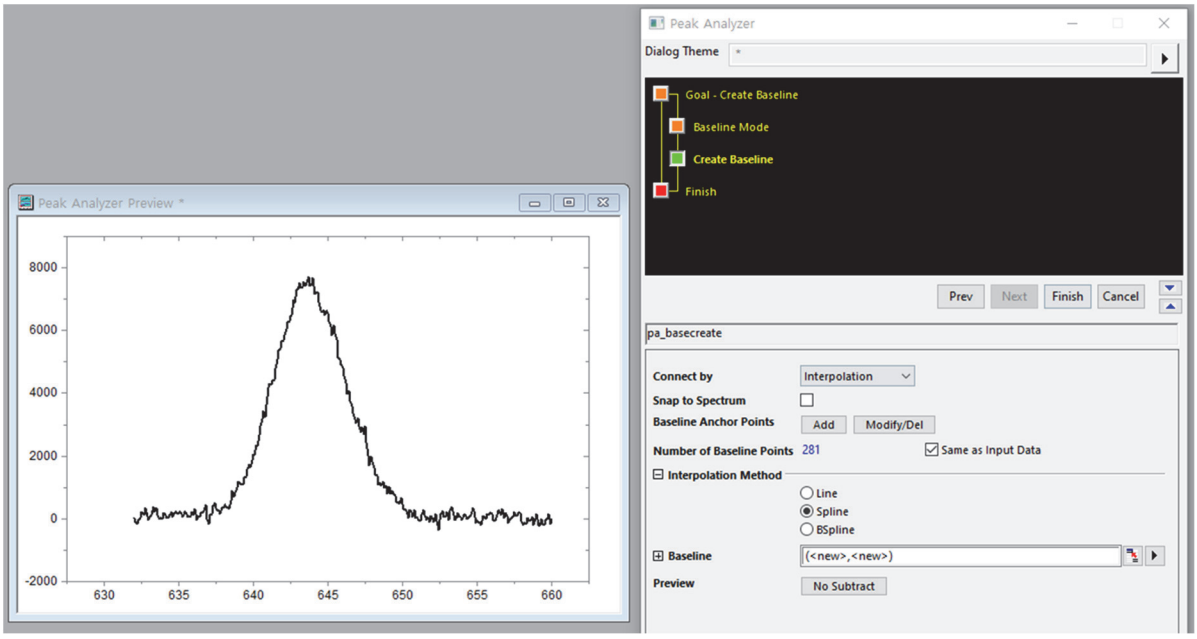

Fig. 4. Refined data after subtracting the chosen baseline from the raw data.

■ 해당 설정을 마치면 그림 6과 같이 창 하단 'Fit Curve'에서 추가적인 peak들이 생성된 것을 확인할 수 있다. 검정색 그래프는 실험 측정 데이터를, 다른 색의 그래프들은 각각 시뮬레이션을 통해 계산되는 peak 1, 2, 3과 그 합계를 나타낸다.

■ 해당 fitting은 Voigt 함수를 기반으로 하기 때문에, 앞서 설정한 3개 peak의 중심값을 입력해 주어야 한다(그림 7).

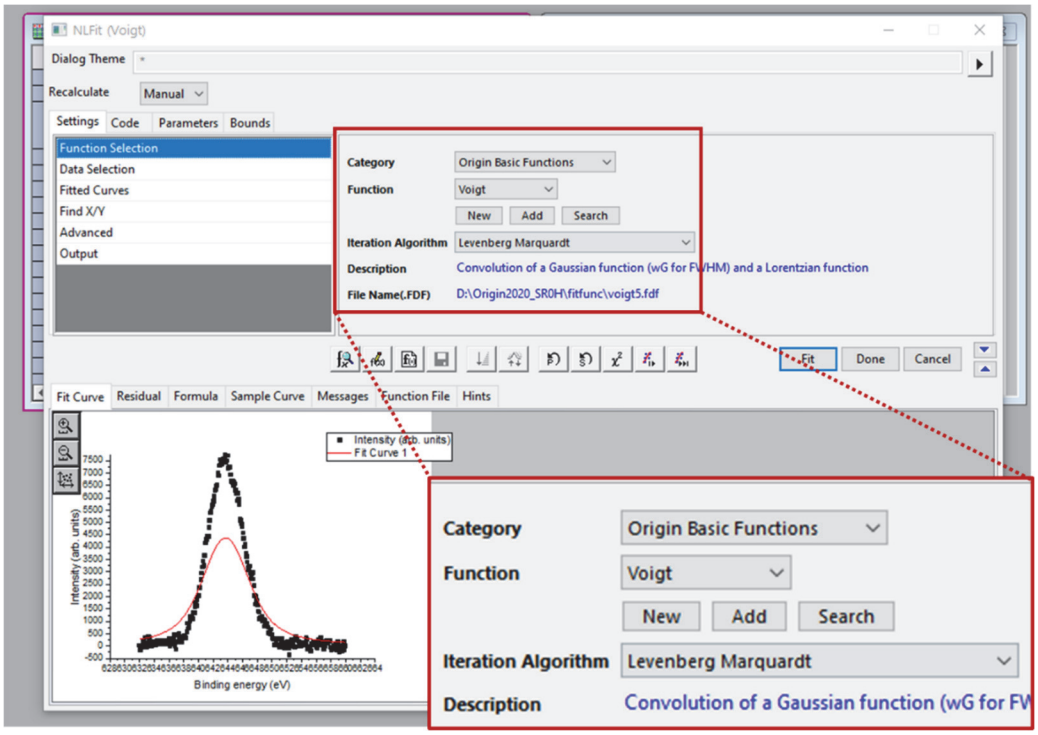

Fig. 5. Function set-up in the function selection menu.

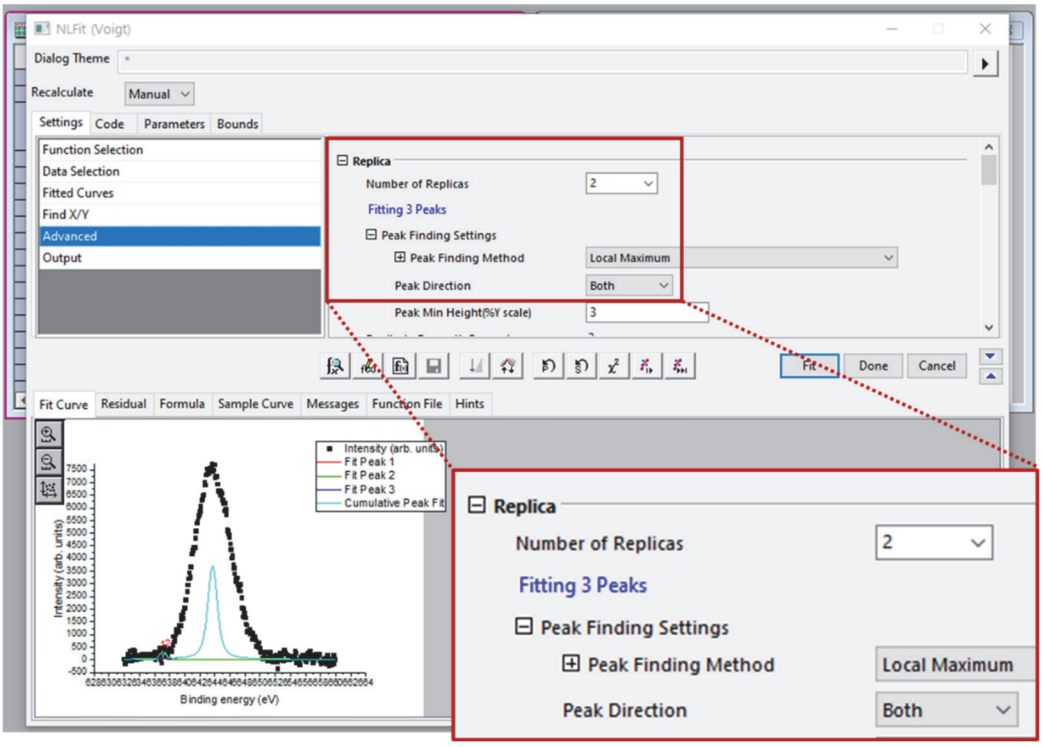

Fig. 6. Replica set-up in the advanced menu, where one can select the number of peaks to be fitted.

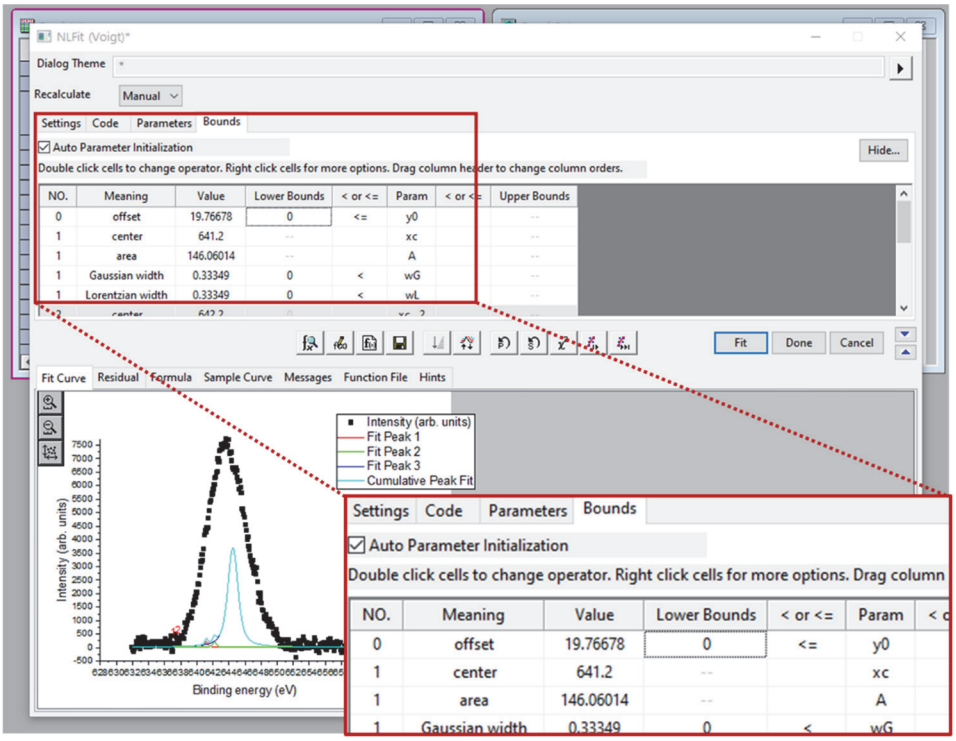

Fig. 7. Peak center and bounds set-up by considering the binding energy of each ion's energy state.

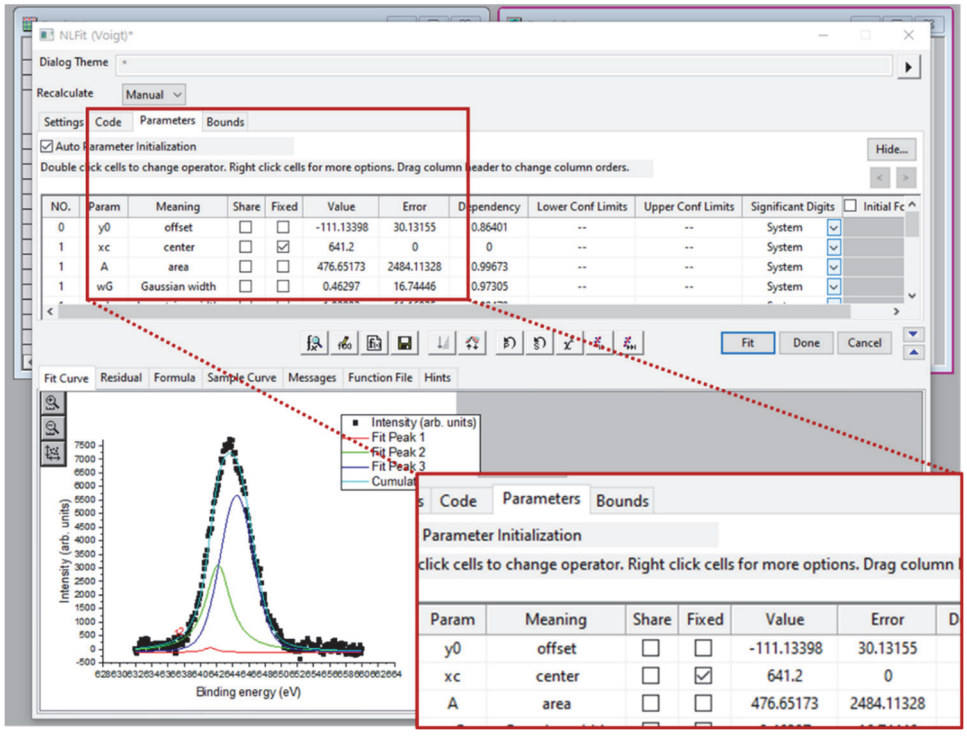

Fig. 8. Actual fitting process by varying parameters.

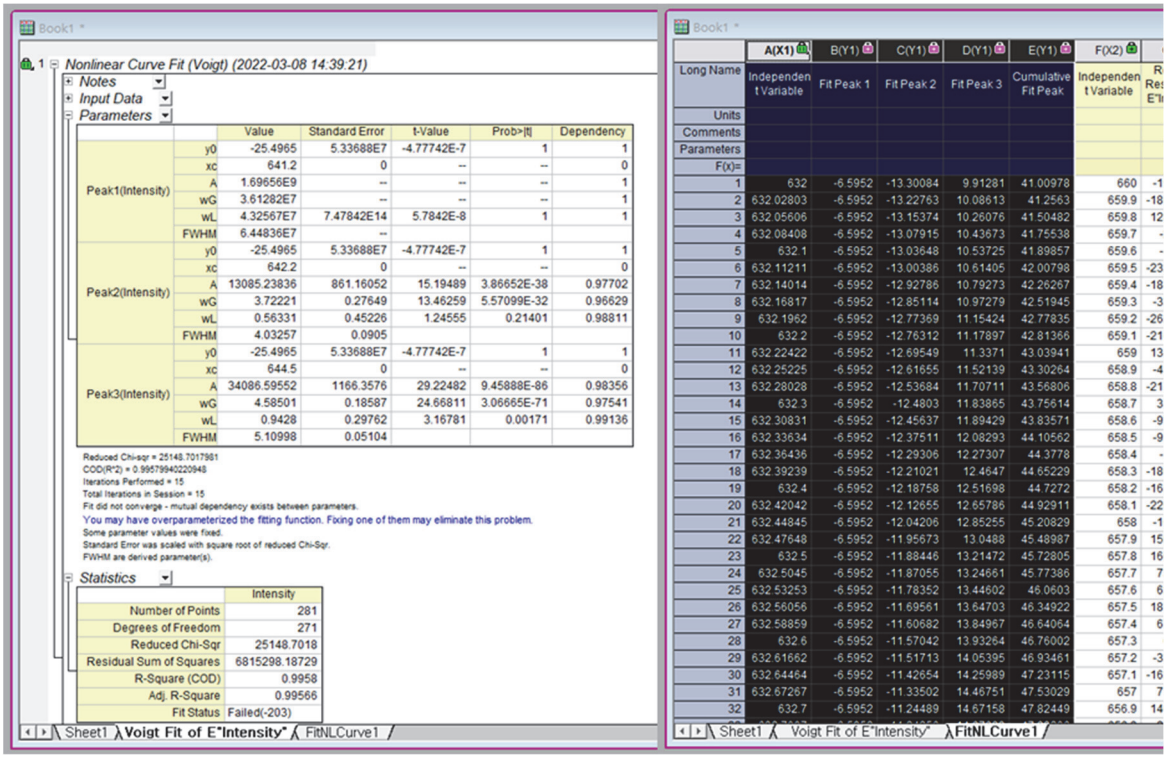

Fig. 9. Output data for the currently evaluated X-ray photoelectron spectroscopy (XPS).

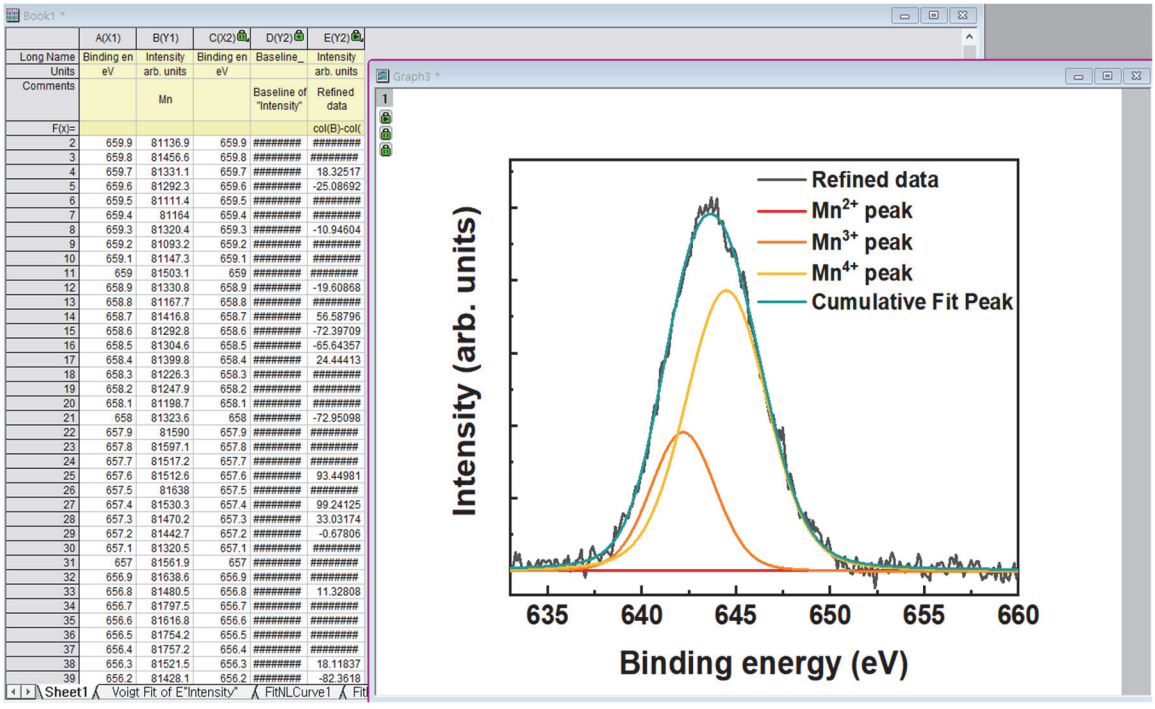

Fig. 10. Collection of the refined and fitted data.

- 각 peak의 시뮬레이션 변수는 왼쪽 숫자를 통해 구별 가능하다. 이는 'Parameters' 또는 'Bounds' 메뉴에서 설정이 가능하며, 중심값은 각 산화수에 따른 타깃 이온(본 report에서는 Mn)의 binding energy를 고려하여 입력한다.
- 모든 fitting은 peak 모양과 실제 데이터의 차이가 최소가 되는 결과치를 도출하는 것을 기본으로 한다. 즉, 실데이터와 fit 곡선과의 차이값이 국소적 최솟값(극소값, local minimum)을 만족하면 fitting이 완료된 것으로 간주한다. 그렇기 때문에 특히 다변수 multi-peak fitting의 경우 주의가 요구되는데, 이는 local minima가 무수히 많이 존재할 수 있기 때문이다.
- 실제로 단순히 기계적 fitting을 진행하는 경우, 거꾸로 뒤집힌 peak이 나타나기도 하고, 변수가 물리적으로 불가능한 값에 setting되기도 한다.
- 따라서, 잘못된 fitting 결과를 피하기 위해 fitting 과정에서 가능한 한 모든 변수가 물리적으로 의미 있는 범위에서 지정될 수 있도록 최저/최고점을 지정해 주어야 한다. 이는 'Bounds' 메뉴에서 설정이 가능하며, 개별 peak의 bounds 영역을 baseline과 refined 데이터의 intensity 크기를 고려하여 설정한다.
- 관련 변수들의 입력이 끝난 뒤, 본 예시에서와 같이 XPS 분석의 경우, 각 산화수에 대한 에너지 준위가 이미 알려져 있기 때문에 문헌을 참고하여 'Parameters' 메뉴에서 각 peak의 center 값을 고정해 주는 과정이 필요하다 (그림 8).
- 즉, 중심값은 산화수에 따른 이온의 고유 물성을 반영하여 입력해준 수치이기 때문에, fitting 과정 중 변동이 생기면 안 되는 값이기 때문이다. 설정이 끝난 후, 'Simplex'를 통해 시뮬레이션을 진행할 수 있고, 'Initialize Parameters'를 통해 fitting 결과가 비정상적으로 나타나는 경우 초기화가 가능하다.
- 이러한 과정은 창 하단의 'Fit Curve'를 참고하고 앞서 'Parameters'와 'Bounds' 메뉴에서 입력한 변수들을 조정하여, 가장 최적의 조건을 찾을 때까지 진행한다. 'Simplex'-'Initialize Parameters'-변수 제어 과정이 끝났다면, 'Fit' 버튼을 눌러 시뮬레이션을 마무리한다.
- 시뮬레이션을 마무리하면 output 데이터 sheet이 생성되는데, 시뮬레이션을 통해 얻어진 각각의 peak 데이터뿐만 아니라, fitting 시 고려된 변수들도 확인 가능하다 (그림 9).
- 해당 변수들을 활용하면 'Tools'-'Fitting Function Organizer' 기능을 통해 x축 범위 재설정, peak 영역 재설정 등을 통해 원하는 형태의 data로 출력이 가능하다.
- 얻어진 시뮬레이션 데이터를 이전 refined 데이터와 함께 정리하면 그림 10과 같은 ready-to-publish 수준의 그래프를 얻을 수 있다.
- XPS 분석에서 가장 중요한 정보인 각각의 산화수를 가진 이온의 양적 비율은 결과 보고서에 출력되는 각 peak의 'Integrated Area' 또는 fitting 결과로 주어진 각 peak에 대한 적분값의 비율을 통해 얻을 수 있다.
- 본 report에서 소개한 방법을 숙지한다면, X-ray absorption spectroscopy 등 다른 peak profile을 fitting하는 데도 도움이 될 수 있을 것이라 기대된다.

ORCID

Wook Jo https://orcid.org/0000-0002-7726-3154

감사의 글

본 연구는 과학기술정보통신부가 지원하는 한국연구재단의 해외우수연구기관 유치사업 (2017K1A4A3015437)의 연구비 지원을 받아 수행되었습니다.

REFERENCES

[1] G. H. Major, N. Fairley, P.M.A. Sherwood, M. R. Linford, J. Terry, V. Fernandez, and K. Artyushkova, *J. Vac. Sci. Technol., A*, **38**, 061203 (2020). [DOI: https://doi.org/10.1116/6.0000377]

[2] C. W. Ahn, C. H. Hong, B. Y. Choi, H. P. Kim, H. S. Han, Y. Hwang, W. Jo, K. Wang, J. F. Li, J. S. Lee, and I. W. Kim, *J. Korean Phys. Soc.*, **68**, 1481 (2016). [DOI: https://doi.org/10.3938/jkps.68.1481]

[3] A. S. Kim and C. H. Sin, *Vac. Mag.*, **3**, 24 (2016).

[4] W.I.F. David, *J. Appl. Crys.*, **19**, 63 (1986). [DOI: https://doi.org/10.1107/S0021889886089999]