


열적 활성화 반응 데이터 분석 오류 최소화에 대한 제언

이주현¹, 천진성², 이규탁², 조 옥¹ 

¹ 울산과학기술원 신소재공학과

² 삼성전기 MLCC 사업부

초록: 연구를 통해 산출하는 실험 데이터의 가치는 그 데이터를 합리적 가설을 바탕으로 정확하게 분석했을 때 빛난다. 재료 분야의 연구자들이 접하는 실험 데이터의 상당수가 열적 활성화 과정을 통해 발현되는 물리 화학적 특성 값이다. 하지만 열적 활성화 과정에 대한 오해로 인해 잘못된 분석 결과를 도출하고 이를 통해 잘못된 결론을 보고하는 경우가 많다. 본 논문에서는 열적 활성화 과정의 의미를 다시 한번 살펴본 뒤, 실제 잘못 해석된 채로 출간되었던 데이터를 예시로 분석 결과의 유효성을 판단하는 기준에 대한 논의 및 재분석하는 과정을 단계적으로 보여줌으로써 향후 유사한 데이터를 획득했을 시 정확한 분석을 진행할 수 있는 역량을 기르는 데 목적을 두고 있다.

키워드: 열적 활성화 과정, 활성화 에너지, 데이터 분석

How to Avoid Misinterpreting Experimental Data for Thermally Activated Processes

Ju-Hyeon Lee¹, Jinsung Chun², Ku-Tak Lee², and Wook Jo¹

¹ Department of Materials Science and Engineering, Ulsan National Institute of Science and Technology, Ulsan 44919, Korea

² Samsung Electro-Mechanics Co., Ltd, Suwon 16674, Korea

(Received April 11, 2023; Accepted April 11, 2023)

Abstract: The value of experimentally obtained data becomes highest when they are properly analyzed based on valid logics. Many physical and chemical properties such as electrical and magnetic properties, chemical reaction rates, etc. are known to be thermally activated; thus, a proper understanding of thermally-activated processes is of importance. However, there are still a number of papers published with falsely analyzed data. In this contribution, we would like to revisit the meaning of thermally-activated processes, and then reanalyze a data set published misinterpreted. By showing a step-by-step procedure for the reanalysis, we would like to help researchers who may come across such data in the future not to make mistakes in their analysis.

Keywords: Thermally-activated process, Activation energy, Data fitting

✉ Wook Jo; wookjo@unist.ac.kr

Copyright ©2023 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. 서문

- 화학반응에서는 물론 수많은 물리적 현상이 일어나기 위해 특정 크기의 활성화 에너지 장벽(activation energy barrier)을 극복해야 하는 경우가 있는데, 이런 모든 반응과 현상을 열적 활성화 과정(thermally-activated process)이라고 부른다.
- 열적 활성화 과정이라는 용어에서도 알 수 있듯이 원자 또는 분자가 가지는 열 에너지는 온도가 높아질수록 커지기 때문에 열적 활성화 과정을 통한 현상은 온도가 높아지면 빨라지는 특징이 있다.
- 역으로 얘기하면 온도가 높아짐에 따라 가속 및 활성화도가 높아지는 많은 물리/화학적 현상은 열적 활성화 과정으로 해석할 수 있다는 의미이다.
- 열적 활성화 과정에 대한 수학적 모델은 스웨덴 출신 물리화학자인 Svante A. Arrhenius (1859~1927)가 경험식으로 제시하였고, 후에 Boltzmann 분포를 따르는 계(system)에 범용적으로 적용 가능성이 확인되었다.
- Arrhenius의 경험식은 특정 반응 속도 k 를 다음과 같이 표현한다.

$$k = k_0 \cdot \exp\left(-\frac{E_A}{k_B T}\right) \quad (1)$$

- 여기서 k_0 , E_A 및 k_B 는 각각 온도에 무관한 선지수(preexponential factor), 활성화 에너지(activation energy) 및 볼츠만 상수(Boltzmann constant, $\sim 8.62 \times 10^{-5}$ eV/K)이다.
- 선지수의 경우, 온도에 무관한 재료 고유의 값으로 수식에서 알 수 있듯이 온도가 무한대에 이르렀을 때의 속도로 화학반응에서는 원자 또는 분자 간 유효 충돌 횟수와 관련되어 있다.
- 한편, Arrhenius의 경험식을 볼츠만 분포 관점에서 기술하면 아래 식과 같이 표현된다.

$$\frac{n}{N} = \exp\left(\frac{\Delta S}{k_B}\right) \exp\left(-\frac{\Delta H}{k_B T}\right) \quad (2)$$

- 여기서 n 은 전체 N 개의 입자 중 ΔH 만큼의 에너지를 받아 활성화된 입자의 개수를 의미하고, 앞서 Arrhenius 경험식에서의 선지수는 활성화된 입자와 활성화되지 못한 입자들로 구성된 계의 배열 엔트로피(configurational entropy, ΔS) 값으로 주어진다.
- 이 경우 역시 수학적으로 선지수는 계의 온도가 무한대에 이르렀을 때 활성화될 수 있는 최대 입자의 분율을 의미하지만, 실제로는 구성 입자의 진동 주파수가 최대 값에 달하는 Debye 온도가 그 기준 값이 된다.

- 관련된 이론적 배경에 대한 보다 자세한 내용은 시중에서 쉽게 구할 수 있는 다양한 열역학 서적의 통계 열역학 부분을 참조할 수 있기 때문에 본 기고에서는 실험 데이터 해석과 관련된 보다 실제적인 내용을 다루고자 한다.

2. 본문

2.1 x, y축 scale과 fitting 신뢰도의 상관관계

- 그림 1에 제시된 실험 데이터에서 보듯이 특정 물리적 특성이 온도에 따라 비선형적으로 증가하는 경우 해당 물성은 열적 활성화 과정으로 분류할 수 있다.
- 열적 활성화 과정의 실험 데이터를 해석하는 과정에서 직면할 수 있는 다양한 상황을 보다 효과적으로 보여주기 위해 그림 1의 데이터는 의도적으로 참고문헌 [1]에 제시된 완화형 강유전 재료의 유전 특성 데이터를 차용하였고 모든 fitting 및 illustration은 상용 소프트웨어인 OriginPro™를 사용하였음을 밝힌다.
- 주지하듯이 그림 1과 같은 데이터는 아래와 같이 Arrhenius 식의 양변에 \ln 를 취한 후 온도의 역수를 변수로 잡으면 직선으로 표현되기 때문에 누구나 쉽게 fitting이 가능하다.

$$\ln k = \ln k_0 - \frac{E_A}{k_B} \cdot \frac{1}{T} \Rightarrow y = a + bx \quad (3)$$

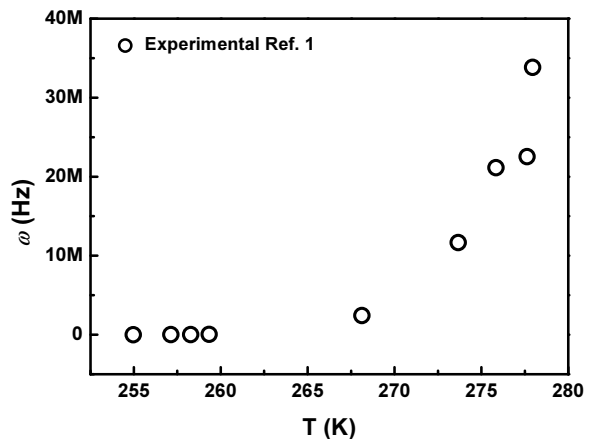


Fig. 1. The measurement frequency for the frequency-dependent dielectric maximum in the prototypical relaxor ferroelectric $\text{Pb}(\text{Mg}_{1/3}\text{Nb}_{2/3})\text{O}_3$ ceramic as a function of temperature [1] (a nonlinear increase in the frequency values with increasing temperature suggests a thermally-activated process as the underlying mechanism).

- 즉 직선 fitting 후 y 절편과 기울기를 통해 선지수 값과 활성화 에너지를 도출할 수 있다는 의미이다.
- 하지만 지금은 개인 컴퓨터마저도 성능이 우수하기 때문에 혹자는 log-log plot을 통한 직선 fitting을 굳이 하지 않더라도 지수함수 식 그 자체로 fitting을 진행해도 무방하다고 생각할 수도 있다.
- 실제로, 대부분의 비선형 데이터 fitting이 해당 현상의 물리적 모델에 기반한 비선형 함수를 통해 진행되고 있음은 주지하는 사실이다.
- 따라서 지수적으로 변하는 실험 데이터의 fitting을 있는 그대로의 비선형 모델을 통해 fitting하는 것과 log-log plot을 적용하는 것 사이에 어떤 차이가 있는지를 명확히 확인하고 넘어갈 필요가 있다.
- 그림 2(a)는 앞서 논의한 Arrhenius 식을 그대로 적용해서 그림 1의 데이터를 fitting한 결과이다. 눈으로 보기에 실험 데이터를 잘 모사하고 있는 듯 보인다. 실제로 이 fitting의 경우 조정 후 상관관계(adjusted R-squared) 값은 0.95599로 실험 데이터와 fitting 모델 간의 높은 상관성을 보여준다.
- 하지만 실험 데이터와 fitted 데이터를 log-log space로 옮기면 우측 그림에서 보듯이 저온부의 데이터를 전

혀 설명하지 못함을 알 수 있다.

- 한편, 그림 2(b)에 제시한 것처럼 실험 데이터를 log-log space로 옮겨서 linear fitting을 진행할 경우, 눈으로 보기에는 fitting 결과가 실험 데이터를 제대로 모사하지 못하는 것처럼 보이지만, 우측 그림에서와 같이 실측 scale로 옮겼을 때 그림 1(a) 못지않은 정확도를 보임을 알 수 있다. 이 경우 조정 후 상관관계는 0.9605로 그림 1(a) fitting 결과와 동일한 수준이다.
- 많은 경우 조정 후 상관관계 값이 1에 가깝다는 것을 fitting 결과가 실험 데이터를 잘 설명한다는 주장에 대한 객관적 근거로 사용하고 있지만, 제시한 예에서 보는 것처럼 단순히 조정 후 상관관계 값을 기반으로 fitting의 유효성을 주장하는 것은 상당히 위험할 수 있음을 주지할 필요가 있다.
- 비선형적 데이터의 fitting 시 생길 수 있는 이런 문제를 보다 명확하게 이해하기 위해서는 최소 자승법(least square method)에 기반한 fitting algorithm을 이해할 필요가 있다.
- 그림 3(a)은 실험 데이터의 독립 변수와 종속 변수 간에 선형적 관계가 있다는 가정 하에 수행된 선형 fitting이다. 최소 자승법이란 주지하다시피, fitting 대상이 되

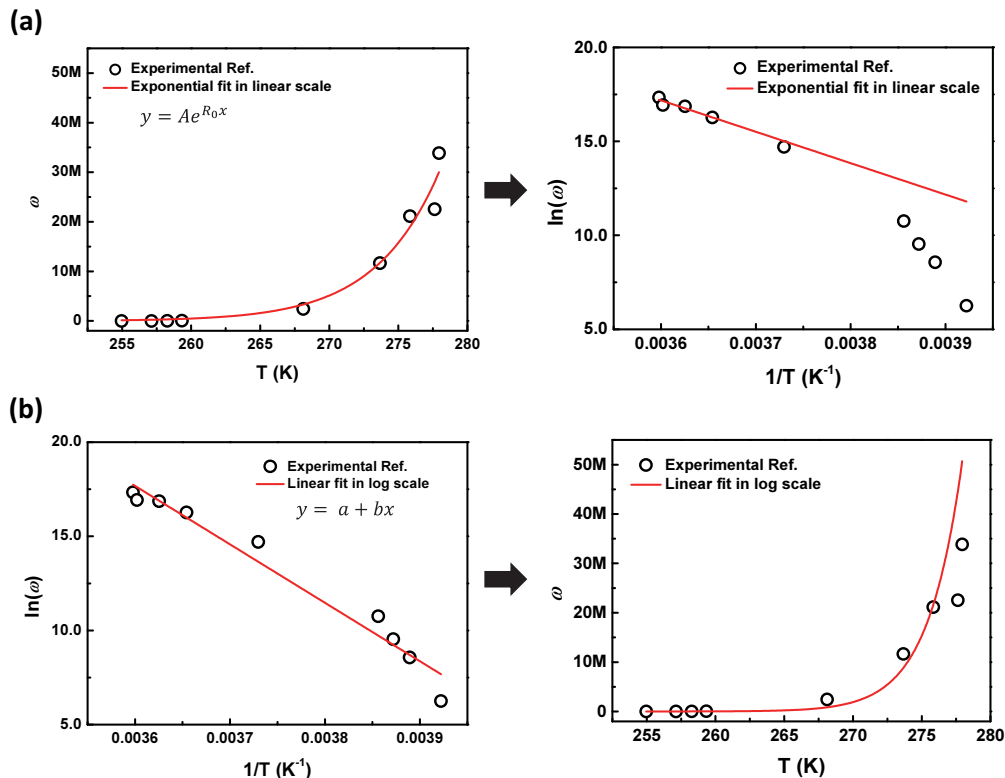


Fig. 2. A comparative illustration of fitting accuracy (a) by a non-linear curve fitting directly on raw data and (b) by a linear fitting on data replotted in a log-log space.

는 데이터와 fitting 함수 간 거리(e_i)가 fitting의 오차에 해당하므로 그 거리의 합(보다 엄밀하게 말하면, $\sum_i e_i^2$)이 최소가 되는 함수를 찾는 과정이다.

- 한편, 그림 3(b)에 제시된 바와 같이 fitting 함수가 비선형성을 가지는 경우, 같은 수준의 x 변화량에 대해 y 변화량의 변화가 위치마다 다르게 나타난다. 예를 들어 곡률이 큰 파란색으로 표시된 부분은 빨간색으로 표시된 부분과 동일한 x 값의 변화에 대해서도 상당히 큰 y 값 변화를 보인다. 즉 최소 자승법의 관점에서는 파란색 부분을 중심으로 오차를 줄이는 것이 조정 후 상관관계 값을 줄이는 데 더 효율적이라는 의미이다.
- 따라서 그림 2에서 실측 값을 지수 함수로 fitting한 결과에서 본 바와 같이 fitting 결과를 log-log plot으로 바꾸었을 때 저온부 데이터 fitting에 큰 오류가 발생하는 것이다.

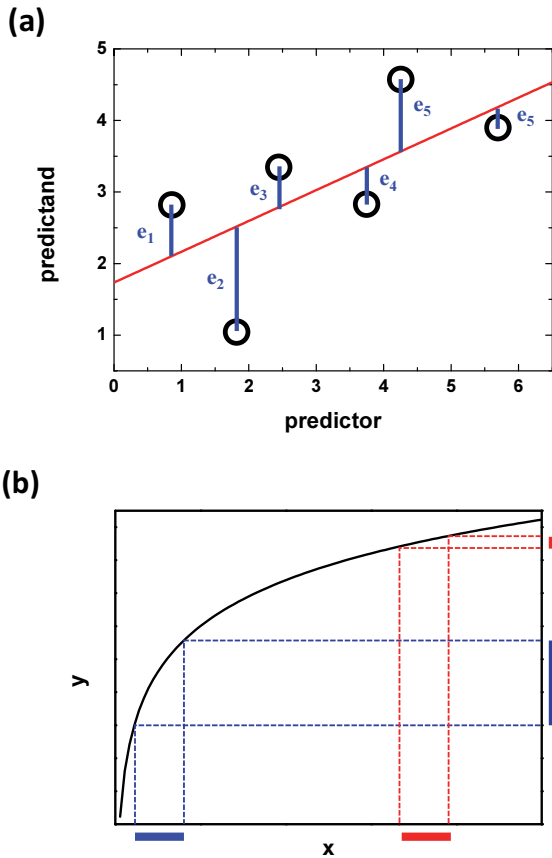


Fig. 3. (a) An example depicting how the commonly practiced least square method works and (b) a schematic illustrating why the scale of x and y axis should be properly chosen.

2.2 Log-log plot에서의 비선형 데이터 해석

2.2.1 통상적 접근

- 그림 2의 데이터를 다시 한번 들여다보면 log-log space에서도 데이터 분포가 확연히 선형성을 벗어남을 알 수 있다.
- 이렇게 열적 활성화 과정에 기인하는 물리현상 데이터를 log-log space에 표시했을 때 선형성을 벗어나는 경우 통상적으로 주 열적 활성화 메커니즘이 저온부와 고온부에서 변하는 것으로 가정하고 두 개의 직선으로 fitting을 진행한다.
- 그림 4는 앞서 논의한 그림 2(b)의 데이터가 온도에 따라 두 가지 다른 메커니즘에 의해 발현된 것이라는 전제 하에 두 개의 직선으로 fitting하는 과정을 보여준다. 실제로 사용한 데이터의 원본 논문 [2]에서도 두 개의 직선을 통해 분석하였다.
- 예전에 컴퓨터 소프트웨어가 다양하지 않던 시절에는 어떤 데이터가 어떤 선에 포함되어야 할지에 대한 명확한 기준이 없어 대략적인 눈대중으로 분석했기 때문에 분석자 스스로도 분석에 대한 확신이 부족했던 반면, 제시된 바와 같이 onset point 분석법을 활용하면 논란의 여지가 없는 분석이 가능하다.
- 그림 4에 제시된 onset point 분석법은 Origin™에서 제공하는 onset 분석툴 중 two-point 분석법으로 분석자가 지정한 영역의 양끝단의 점에서 접선의 기울기를 통해 두 개의 직선을 분할해 주는 방법을 사용했다. Two-point 분석법 외 four-point 분석법도 있는데, 자세한 설명한 다음 section에서 기술하고자 한다.
- 그림 4(a)에서 보듯이 onset point 분석법을 활용하면 단일 직선으로 fitting이 되지 않는 데이터의 경우 두 개의 직선을 활용하면 깔끔하게 fitting이 됨을 확인할 수 있다.
- 특히, onset point 분석법을 활용하는 경우, 각 직선의 기울기(E_A)와 교점의 좌표를 도출해 주기 때문에 그림 4(b)에 제시한 것처럼 fitting 함수의 재구성이 용이하다는 장점이 있다.
- 하지만 두 개의 직선을 통한 분석의 경우 외관상 fitting이 잘되었다 하더라도 분석 결과가 합리적인지에 대한 검토가 필요하다. 그림 4의 데이터에 대한 분석 결과에 따르면, 고온부에서 작동하는 열적 활성화 메커니즘의 활성화 에너지가 1.78 eV이고 저온부의 경우 6.03 eV로 제시되는데, 두 경우 모두 활성화 에너지가 지나치게 높다는 점에서 분석의 유효성이 의심되기 때문이다.

- 활성화 에너지 값의 높고 낮음은 해당 온도에서의 열 에너지(thermal energy, $k_B T$)의 크기가 얼마인지를 통해 가늠할 수 있다. 관련 분석이 많은 연구자는 기준 값으로 상온(298 K)에서의 열 에너지가 0.026 eV임을 기억하고 있으면 편리할 것이다.
- 통상적으로 사용하는 벌크(bulk) 세라믹 소재의 유전율은 두께 1 mm, 지름 10 mm 내외 크기의 시료에 RMS (root-mean square) 1 V 정도의 전압을 인가해서 측정한다. 즉 측정을 위해 인가된 전기장에 의해 추가되는 에너지가 아주 미미하다는 점을 감안할 때, 측정 온도대에서의 열 에너지가 현상을 유도하기에는 활성화 에너지가 지나치게 높다는 의미이다.
- 실제로 slope 2로 fitting된 process의 경우 265 K (~8°C) 근방에서 급격하게 활성화되는 것을 알 수 있는데, 인가한 전기 에너지와 그 온도에서의 열 에너지의 조합으로 6.03 eV라는 큰 활성화 에너지를 극복할 수 있다고 하는 것은 논리적으로 불합리하다고 보는 것이 타당하다.

- 뿐만 아니라 그림 4 데이터를 처음 보고한 원문에서도 확인할 수 있듯이 고온부, 즉 열 에너지가 더 높은 영역에서 작동하는 process의 활성화 에너지가 저온부의 활성화 에너지보다 더 낮은 것에 대해 논리적, 합리적 근거를 제시하기 힘들다.
- 사실 이와 같이 log-log plot에서 비선형인 열적 활성화 process 데이터 해석과 관련한 논란은 재료공학, 화학, 및 생명공학 등에서 아주 오래전부터 이슈가 되어 왔는데 관련된 논의에 대해 조금 더 살펴보고자 하는 독자 들은 참고문헌 [3]을 참고하면 도움이 될 것이다.
- 따라서 비단 지금의 예에 국한된 것이 아니라 실제로 많은 열적 활성화 과정이 Arrhenius 식에서 벗어난다는 것이 이미 잘 알려져 있기 때문에 이들 실험 결과를 설명하기 위한 이미 다양한 모델들이 제시되어 있다.
- 대표적인 예로, 반도체의 전기전도의 예를 들 수 있다. Doping에 의한 외인성(extrinsic) 영역에서 진성(intrinsic) 영역으로 바뀔 때 따라 활성화 에너지가 바뀌게 되는데, 이런 경우가 대표적으로 앞서 소개한 바

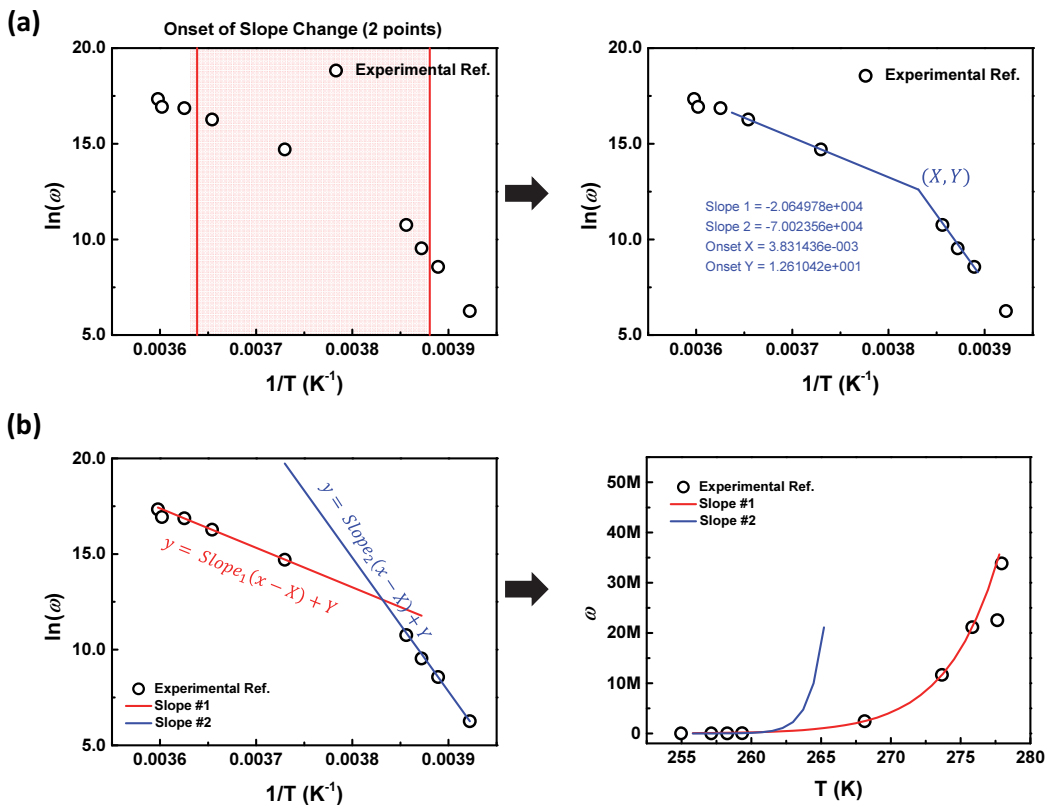


Fig. 4. (a) An illustration showing how to fit a non-linear data by assuming that the process originates from two physically different mechanisms [this type of fitting can be easily done in the Origin™ software simply by installing an add-on provided by the manufacturer (The details of this information are given in the next section)] and (b) a procedure how to construct fitting results graphically from the onset data analyzed in (a).

와 같이 특정 온도를 경계로 두 개의 서로 다른 메커니즘이 경합하는 방식으로 설명된다.

- 이때, doping된 외인성 반도체의 온도에 따른 전기전도도 변화는 그림 4에 제시된 데이터와 달리 저온의 외인성 영역에서 activation energy가 고온의 intrinsic 영역에서 보다 상당히 낮다는 점을 주지할 필요가 있다. 즉 두 개의 다른 process를 전제로 두 개의 직선을 활용하는 경우에는 반드시 고온부의 activation energy가 저온부보다 높은지를 확인할 필요가 있다는 의미이다.
- 저온부와 고온부에서 서로 다른 두 개의 메커니즘이 작동하는 열적 활성화 과정에 대한 분석이 필요한 경우 다음 section에 소개한 방법을 익혀 두면 유용할 것이다.

2.2.2 Onset point 분석

- 과거에는 nonlinear 데이터를 두 개의 직선으로 구분하여 fitting할 때 말 그대로 눈대중으로 두 개의 가상의 선을 그은 다음 각 선에 속하는 데이터를 따로 분리하여 각각의 데이터에 대해 선형회귀법을 적용하여 활성화 에너지 및 필요 정보를 얻었다.

- 하지만 이런 방법은 다분히 주관적이며 오류가 많을 뿐만 아니라 그 과정이 번거롭기 때문에 관련 분석이 많은 분들을 위해 Origin에서 제공하는 module을 소개한다.
- 원래 상전이 온도, 유리 전이 온도 등과 같은 열분석 과정에서 중요한 의미를 지닌 onset point를 간편하게 찾아낼 수 있도록 하는 용도로 origin에서 개발한 module인데, 해당 module의 이름은 OL_Estimate_Onset.OPK이고 <http://www.originlab.com/fileexchange/details.aspx?fid=107>에서 무료로 받을 수 있다.
- OL_Estimate_Onset.OPK module을 실행하면 onset point를 분석하는 두 가지 선택이 있는데, 각각의 예시를 그림 5에 도시하였다.
- 2-point 분석은 선택 영역의 끝점에서의 접선을 그어 각 접선의 기울기 및 두 접선의 교점, 즉 onset point를 알려준다.
- 반면, 4-point 분석은 선택된 각 영역의 끝점을 지나서 두 개의 직선을 구하고, 각 직선의 기울기 및 그 교점을

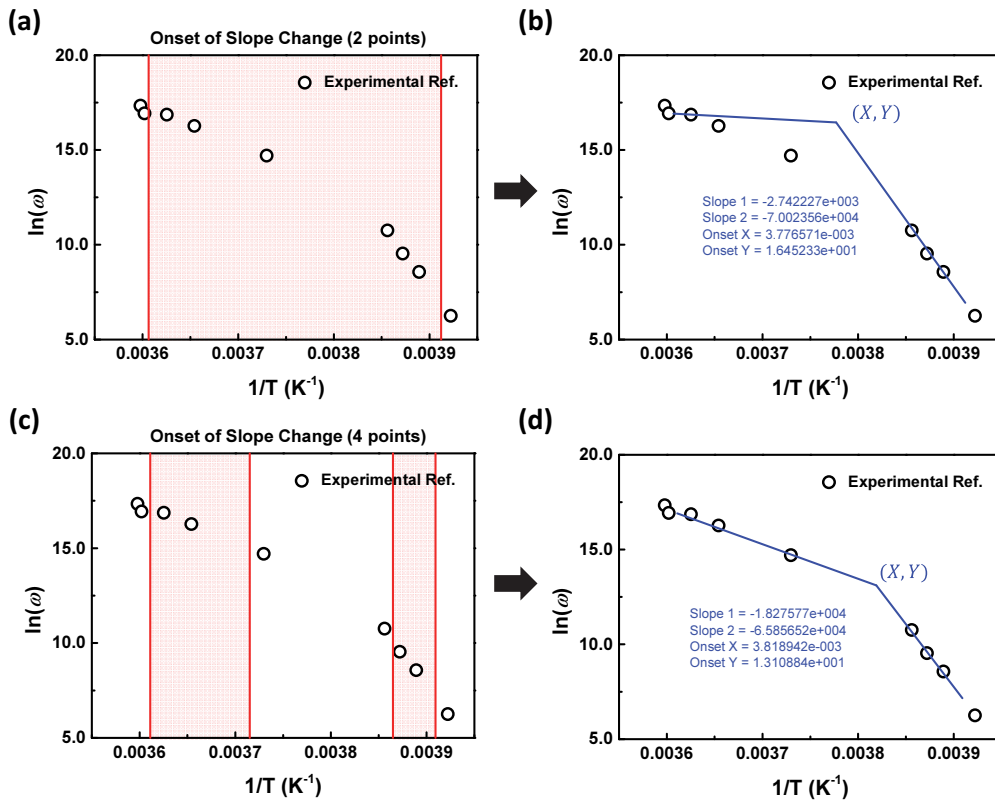


Fig. 5. Comparative illustration explaining two different ways to fit non-linear data with two straight lines using Origin™ software: (a) & (b) 2-point estimation and (c) & (d) 4-point estimation.

찾아준다.

- 어떤 방법이 더 나은 방법인지에 대한 절대적 기준은 없기 때문에, 분석자가 상황에 맞게 보다 적절한 방법을 선택하여 분석을 진행하면 된다.

2.2.3 Log-log plot 상에서 비선형인 데이터 분석의 사례

- 앞선 section에서 그림 1에 제시된 데이터의 경우 고온 영역에서 저온 영역보다 더 낮은 활성화 에너지를 필요로 하는 process가 진행됨을 전제해야 한다는 점만으로도 두 개의 직선을 이용한 fitting은 적용하는 데 무리가 있음을 논의했다.
- 즉 log-log 공간에서 $1/T$ 에 대해 위로 볼록한 형태의 궤적을 그리는 데이터는 원칙적으로 두 개의 직선으로 fitting할 수 없다는 의미이다.
- 이전 section에서 언급한 바와 같이 non-Arrhenius 거동을 보이는 데이터를 fitting하는 모델들이 다양하게 제시되어 있지만, 본 논문에서는 그중 물리적 근거가 명확한 소위 Vogel-Fulcher-Tammann (VF) 식을

이용한 fitting을 소개한다.

- VF 식은 원래 유리전이를 설명하기 위해 도입된 개념으로, 다음과 같이 표현된다.

$$A = A_0 \cdot \exp\left[-\frac{E_A}{k_B \cdot (T - T_0)}\right] \quad (4)$$

- 앞서 소개한 Arrhenius 식과 비교하면 단순히 T 가 $T - T_0$ 으로 바뀌었다는 것을 쉽게 확인할 수 있다. 즉 VF 식에서 $T_0 = 0$ 인 특수한 경우가 우리가 흔히 알고 있는 Arrhenius 식이라는 의미이다.
- 주지하듯이 절대온도 0 K 은 어떠한 열적 활성화 과정도 일어나지 않는 온도이다. 마찬가지로 VF 식에서 T_0 은 VF 식을 따르는 특정 열적 활성화 과정이 더 이상 생길 수 없는 온도로 이해할 수 있다.
- VF 식은 $T \gg T_0$ 의 조건에서 $T - T_0 \approx T$ 로 볼 수 있기 때문에 고온에서는 Arrhenius 식과 근사하지만 온도가 낮아지면서 T_0 에 가까워질수록 Arrhenius 식에서 급격히 벗어남을 쉽게 이해할 수 있다. 즉 Arrhenius 식에 T_0 이라는 parameter 하나를 추가함으로써 쉽게

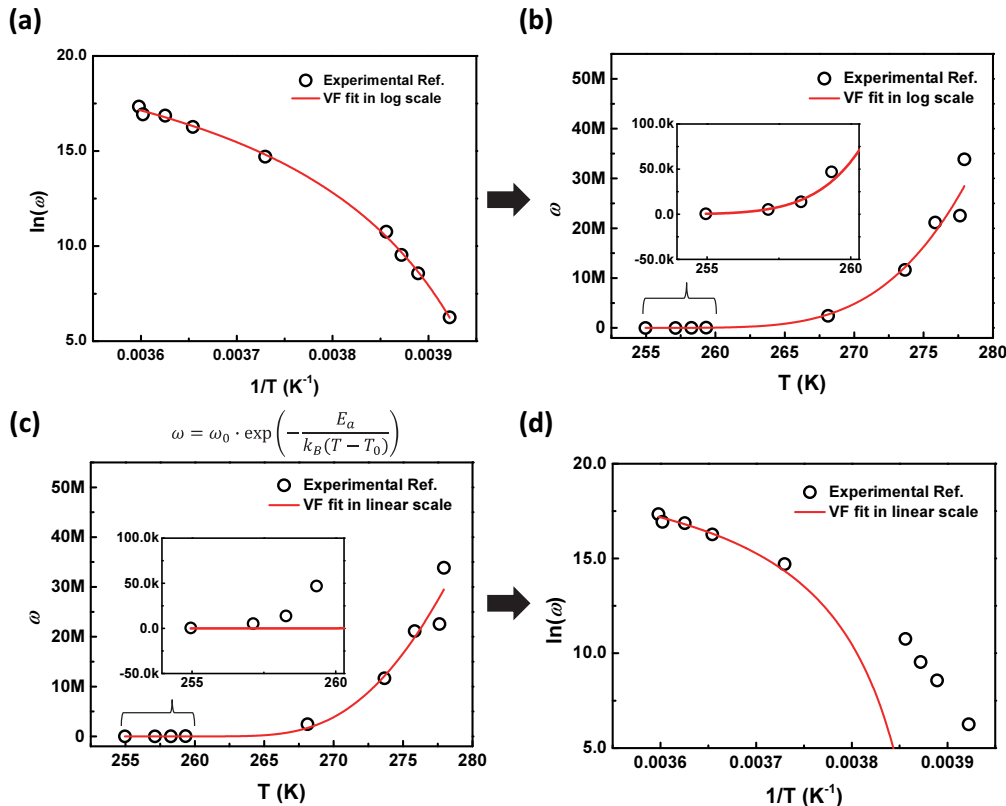


Fig. 6. (a) Fitting the data from Fig. 1 in a log-log space using VF equation (the fitting result can be seen extremely satisfactory both in the log-log space), (b) the original linear scale, and (c) as discussed earlier, fitting exponentially changing data in the original linear scale even with the most appropriate fitting equation can lead to a significant error, which can be best-seen in the corresponding log-log plot.

log-log plot 상에서 위로 볼록한 곡선을 구현할 수 있다는 의미이다.

- 그림 6(a)는 그림 1에 제시한 data를 log-log space에서 VF 식을 사용하여 fitting한 결과이다. Fitting의 정확도가 한눈에 봐도 높음을 알 수 있다. 물론 그림 6(b)에서 보듯이 원래의 선형 공간으로 되돌려도 주어진 데이터 잘 설명하고 있음을 알 수 있다.
- 이때, 활성화 에너지 값이 0.02 eV로 주어지는데, 상온 부근에서의 열 에너지와 비슷한 수준의 값이므로 현재의 열적 활성화 고정을 잘 설명하고 있음을 확인할 수 있다.
- 그림 6(c) 및 (d)에 제시한 것처럼 아무리 정확한 함수를 사용한 fitting이라도 원래의 선형계에서는 최소 자승법의 원리상 오류가 클 수밖에 없음을 주지할 필요가 있다.
- 본 논문에서 중점적으로 다룬 내용은 열적 활성화 과정을 거치는 실험 데이터를 fitting하는 과정에서 경험할 수 있는 오류 상황에 대한 소개 및 정확한 fitting을 위한 제언이었지만, 많은 물리/화학적 현상이 exponential dependence를 따르고 있음을 감안할 때, 열적 활성화 과정에 의한 데이터 외에 수많은 실험 데이터 해석에 광범위하게 적용할 수 있는 접근법임을 주지한다.

ORCID

Wook Jo

<https://orcid.org/0000-0002-7726-3154>

감사의 글

본 연구는 한국산업기술평가관리원이 주관하는 소재부품기술개발사업(20022441)의 지원을 받아 수행되었습니다.

REFERENCES

- [1] 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>]
- [2] V. V. Kirillov and V. A. Isupov, *Ferroelectrics*, **5**, 3 (1973). [DOI: <https://doi.org/10.1080/00150197308235773>]
- [3] J. Kohout, *Molecules*, **26**, 7162 (2021). [DOI: <https://doi.org/10.3390/molecules26237162>]