

Informe de Métodos Numéricos

Tarea 10: Ajustes de Parámetros Experimentales 2.0

Leonardo Leiva

1. Resumen

En este informe se trabajará a partir de un espectro de absorción de una fuente. Se busca modelar esto de dos formas y luego se determinará si son válidas o no y cual de las dos es mejor.

2. Introducción: Marco Teórico

En este informe describe el proceso para aproximar datos experimentales a dos modelos, ver que tan válidos son dichos modelos y ver cual es mejor.

Los datos experimentales a modelar corresponde a mediciones realizadas por espectroscopía. Consiste en analizar la radiación emitida por fuentes en función de la longitud de onda. Algunas características relevantes que se pueden obtener son la intensidad, forma del continuo, líneas de emisión y absorción, que permiten comprender el comportamiento de la fuente. Algunos usos prácticos son entender la temperatura interna o temperatura (de una estrella, por ejemplo), la composición o la cinética de una reacción química o identificar átomos o moléculas mediante sus espectros^[1].

Se analizará un espectro que se compone de una parte continua y una línea de absorción como se puede ver en la figura (1).

La parte continua puede ser modelada mediante una recta, mientras que la línea de absorción debe ser modelada mediante algún método diferente. En teoría esta línea debería muy delgadas (como una delta de Dirac), pero en la práctica son más anchas. La forma de estas dependerá del mecanismo de ensanchamiento.

En este caso se usarán dos posibles modelos: una Gaussiana; y un perfil de Lorentz. Para ambos casos la idea es modelar simultáneamente con la recta.

2.1. Tipos de Modelo I: Gaussiano

Para el caso del modelo Gaussiano se usará una función del tipo:

$$f(x) = Ae^{-\frac{(x-\mu)^2}{s\sigma^2}} \quad (1)$$

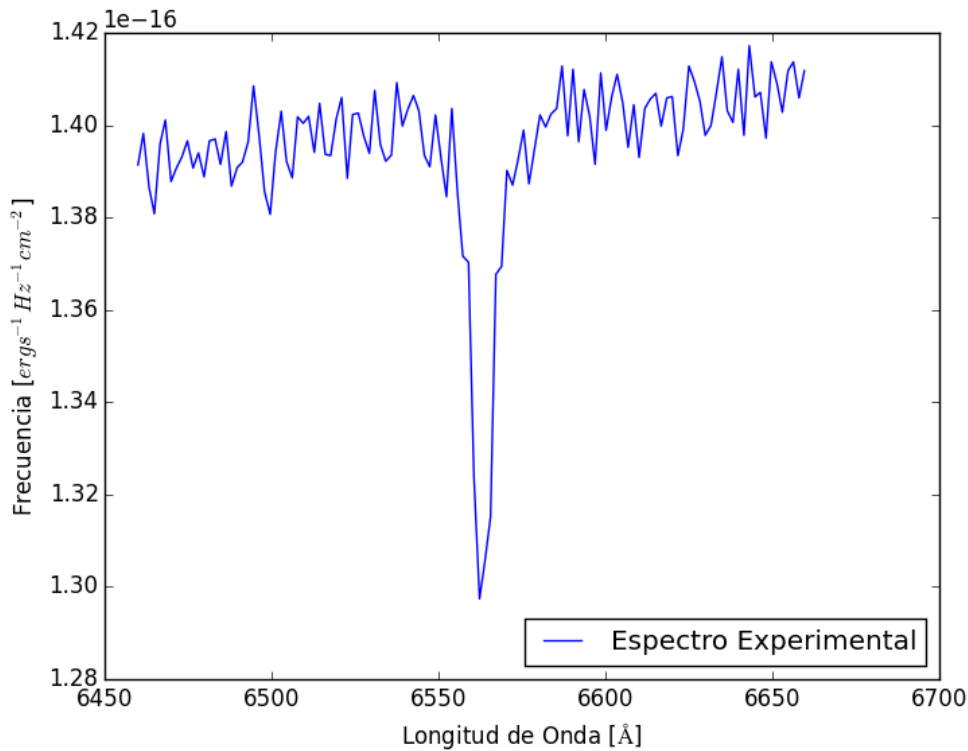


Figura 1: Espectro experimental entregado como dato. Se trabajará con esto en este informe.

Con A la amplitud, μ el centro y σ la dispersión. De esta manera que la frecuencia que buscamos queda modelada por:

$$Frec(\lambda) = a\lambda + b - Ae^{-\frac{(\lambda-\mu)^2}{s\sigma^2}} \quad (2)$$

Donde los primeros términos corresponden a la recta que modela la parte continua. Se buscarán los mejores parámetros a , b , A , μ y σ para los datos experimentales de la figura (1).

2.2. Tipos de Modelo I: Perfil de Lorentz

Para el caso del perfil de Lorentz se usa la distribución de Lorentz (o de Cauchy) que corresponde a:

$$f(x, x_0, \gamma) = \frac{A}{1 + \left(\frac{x-x_0}{\gamma}\right)^2} \quad (3)$$

Donde x_0 indica el punto central y γ es el ancho medio al máximo medio, además se le ha agregado el parámetro A para jugar el rol de la amplitud. La distribución original usa $A = 1$. Son similares a los roles de μ , A y σ para el caso de la Gaussiana. De esta forma, la frecuencia queda modelada como:

$$Frec(\lambda) = a\lambda + b - \frac{A}{1 + \left(\frac{\lambda-x_0}{\gamma}\right)^2} \quad (4)$$

Se buscará los parámetros mencionados que optimicen los datos experimentales.

2.3. Búsqueda de Parámetros

La búsqueda de parámetros que mejor se asemejan a los de los datos experimentales se pueden obtener mediante el algoritmo de Levenberg-Marquardt. Dicho algoritmo fue descrito en el informe anterior, por lo que se omitirá en este caso, pero se puede visitar la referencia [2].

2.4. Prueba de Kolmogorov-Smirnov

Dado que no se sabe con certeza cual de los dos modelos anteriores es el que origina la línea de absorción, resulta importante ver cual de los dos es el mejor. Esto permitiría determinar el origen físico del ensanchamiento. Para saber que tan cierto son los modelos se usa la prueba de Kolmogorov-Smirnov.

Sirve para comparar si dos muestras vienen de la misma distribución. Usualmente se comparan dos muestras empíricas. Se compara a través de las funciones de densidad acumulada. Se calcula la distancia máxima vertical entre la función de densidad de probabilidad de los datos y del modelo D_n . Con esto y el número de datos se puede tener el nivel de confianza del modelo. Mediante la función 'kstest' de scipy se puede obtener D_n crítico y el grado de confianza.

Se explicará de mejor manera lo anterior. Se define H_0 como "Los datos analizados siguen una distribución propuesta". Se define D_n como:

$$D_n = \max_i |F_n(x_i) - F_0(x_i)| \quad (5)$$

Donde x_i es el valor observado de la muestra ordenada, $F_n(x_i)$ es un estimador de la probabilidad de observar un valor menor o igual a x_i y $F_0(x_i)$ es la probabilidad de valores menores o iguales a x_i cuando la distribución de los datos es la que se propone. D_n entonces es la máxima diferencia absoluta entre las funciones de densidad acumulada. Si los datos se parecen a la función modelo se tendrá que D_n es pequeño.

Se aceptará H_0 si $D_n \leq D_n(\text{critico})$. Si no, se rechaza. Se define el D_n crítico a partir de α tal que:

$$P(\text{rechazar } H_0 \mid H_0 \text{ cierto}) = P(D_n > D_n(\text{critico}) \mid \text{Tiene dist. propuesta}) = \alpha \quad (6)$$

Usualmente se toma $\alpha = 0,05$.

3. Metodología

3.1. Modelo 1: Gauss

Para el modelo de Gauss se planteó el modelo de acuerdo a (2) con la función de distribución normal de scipy. Luego, de acuerdo a la forma de los datos experimentales (figura 1) se hicieron algunas adivinanzas sobre los parámetros. La idea es que estuvieran

suficientemente cerca para no converger a un mínimo diferente al que se necesita, pero se verifica que no se necesita tanta precisión en la adivinanza más que la que se puede observar de la figura.

Se escogió $a = 10^{-16}$, a la misma escala que los datos en el eje y , $b = 0$ por el tipo de inclinación de la parte continua de los datos, $A = 10^{-16}$ por la misma razón que a y $\sigma = 1$ porque no se tenía demasiada claridad al respecto y era el parámetro más difícil de adivinar. De cualquier manera no resulta ser demasiado importante.

Con los parámetros iniciales fijos (cabe mencionar que no fueron los primeros parámetros elegidos) se realiza la optimización mediante la rutina 'leastsq' de scipy. Se definieron varias funciones para tener claridad de todos los procesos que se llevan a cabo.

3.2. Modelo 2: Perfil de Lorentz

Para implementar el modelo (4) se usa la distribución de Cauchy implementada por scipy. Aparte de esto, el procedimiento es el mismo en comparación con Gauss. Los parámetros iniciales también se escogen de la misma manera al no haber una razón de peso para cambiarlos (busca modelar los mismos datos y los parámetros son de similares características). Ambos modelos se implementaron en el archivo 'parte1.py'.

3.3. Prueba de Kolmogorov-Smirnov

Se realizó esta prueba en base al demo subido por el profesor^[3]. En primer lugar se ordenan los datos en y tanto experimentales como del modelo. Se define la función de probabilidad acumulada en base a esto. Dado que los datos son discretos se obtiene una función escalonada. El código considera esto para después graficarla. La diferencia máxima D_n y el grado de confianza se pueden calcular a mano, pero se prefiere usar el algoritmo implementado en scipy: 'kstest'. Luego se realiza un gráfico para la función de densidad de probabilidad acumulada. Este test se implementó en el archivo 'parte2.py'.

4. Resultados y Análisis

4.1. Parámetros Obtenidos

Las figuras (2) y (3) contienen los datos experimentales y un modelo en línea continua: la (2) tiene el modelo de Gauss y la (3) el de Lorentz.

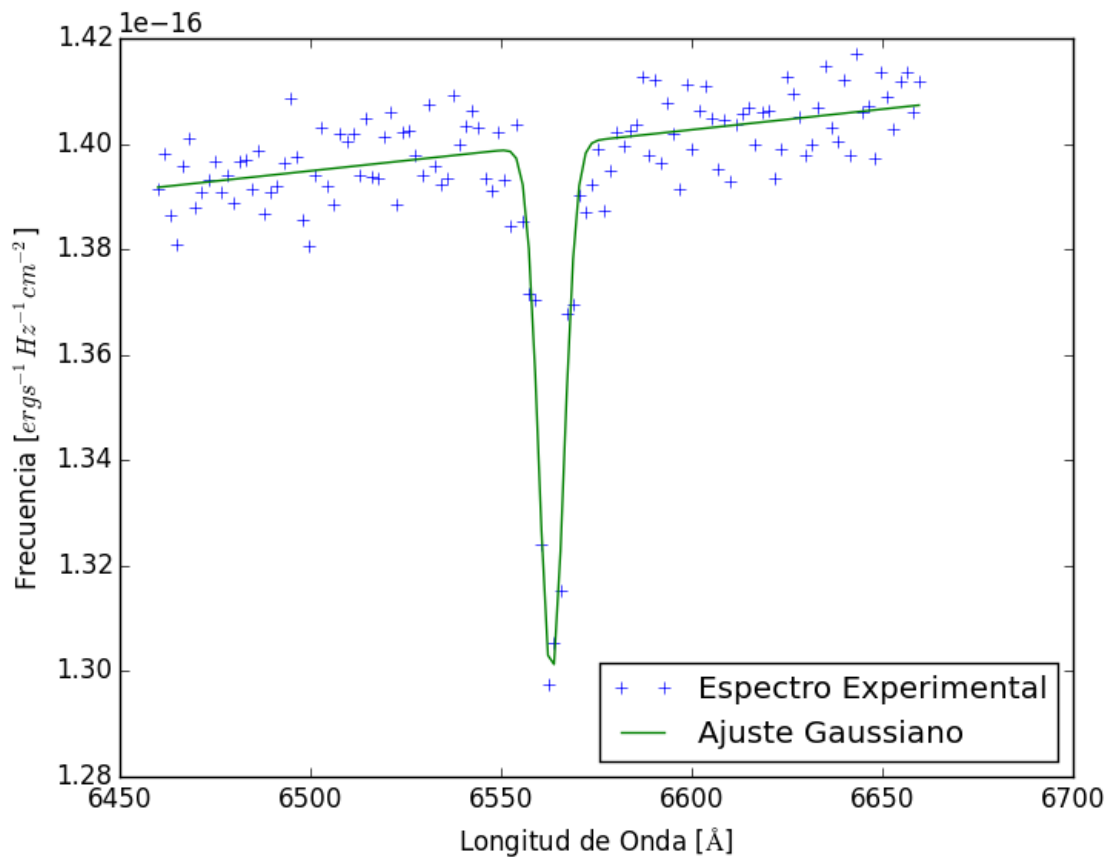


Figura 2: Datos experimentales en azul y ajuste según los parámetros del cuadro (1) para el modelo de Gauss.

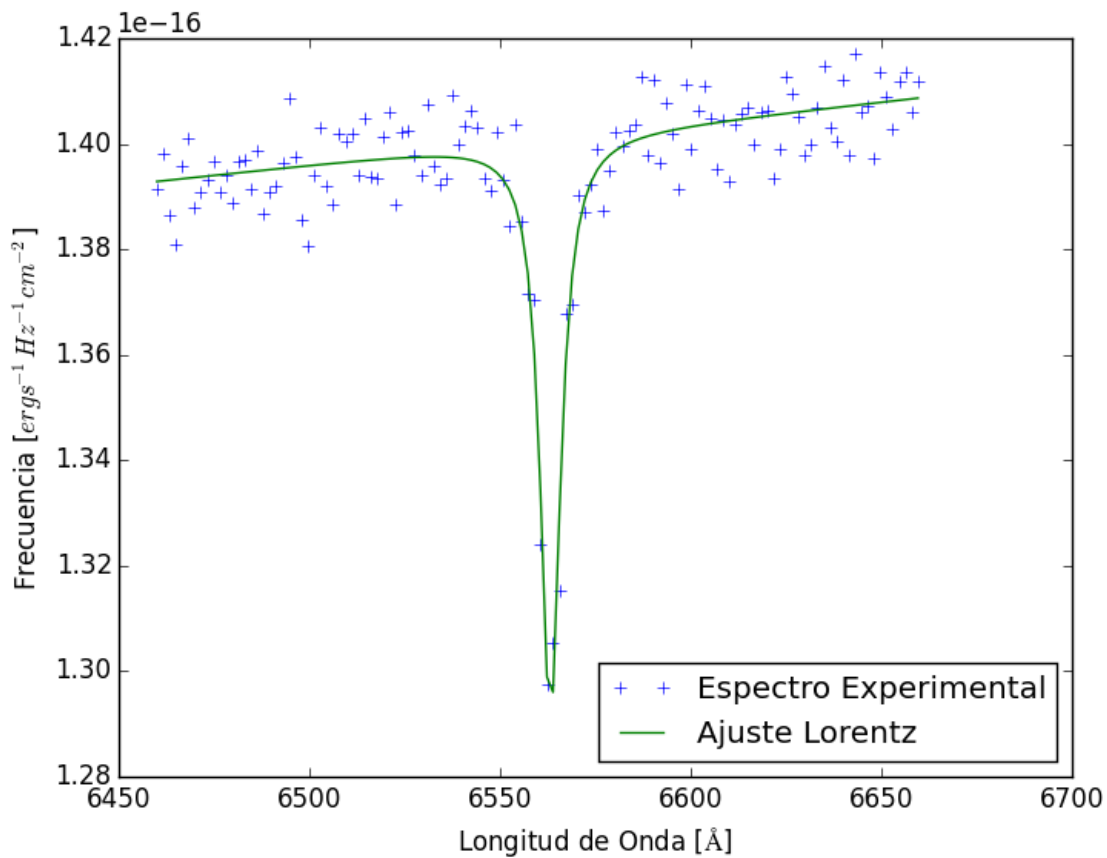


Figura 3: Datos experimentales en azul y ajuste según los parámetros del cuadro (1) para el modelo de Lorentz.

Los parámetros obtenidos con los cuales se graficaron las figuras (2) y (3) se encuentran en la figura 1.

Cuadro 1: Parámetros obtenidos con el Algoritmo de Levenberg-Marquardt

Parámetros	Pendiente (a)	Coef Posición (b)	Amplitud (A)	μ	σ	χ^2
Gauss	$7,8 * 10^{-21}$	$8,87 * 10^{-17}$	$8,22 * 10^{-17}$	6563,2	3,25	$5,2 * 10^{-35}$
Lorentz	$7,92 * 10^{-21}$	$8,81 * 10^{-17}$	$1,1 * 10^{-16}$	6531,1	3,21	$9,62 * 10^{-35}$

Si bien los últimos 2 parámetros no representan exactamente lo mismo, vale la pena compararlos. Se observa que los parámetros son bastante similares: para el caso de la pendiente el orden de magnitud de la diferencia es muy pequeña con respecto al orden de los datos. Se observa además que si bien la adivinanza para este parámetro estaba bien en el sentido de que era pequeña, se obtiene que se requería una pendiente aún menor. Para el caso del coeficiente de posición no tiene demasiada relevancia en el modelo ya que el eje y está muy lejos de los datos y es muy cercano a 0, al igual que la adivinanza. En las figuras se observa que en la parte continua los dos modelos difieren muy poco entre ellos producto

de los parámetros obtenidos, lo cual era esperable.

Para la amplitud se observa una diferencia perceptible en los valores que se refleja muy bien en los gráficos: el modelo de Lorentz alcanza valores más alejados del continuo en la línea de absorción. Esto puede tener suma importancia a la hora del análisis, ya que precisamente se busca estudiar la absorción de ciertas longitudes de onda para la radiación medida. Se cree importante que el modelo deba reflejar la absorción de mejor manera. También se observa que los valores en la zona de absorción quedan mucho más cercanos en Lorentz que en Gauss. De cualquier manera, respecto a estas últimas observaciones se necesitan mas datos para concluir que un modelo sea mejor que otro, pero pueden dar un indicio del resultado óptimo (destacar que se refiere a este criterio. Más adelante se analizará la cercanía de los modelos de acuerdo a los datos que se tiene).

Para los últimos dos parámetros no hay mucho que concluir ya que dependen mucho de la forma que tenga la función. Si se puede notar que existe una diferencia no despreciable en el centro μ que debe ser considerada a la hora de escoger un modelo con respecto a otro porque significaría fijar la longitud en donde ocurre la absorción. En este sentido resulta necesario preguntarse si acaso solo uno de los modelos tiene el resultado real y no podría ser una mezcla de ambos (de una forma similar a lo que se hizo en la tarea anterior para modelar la constante de Hubble, escogiendo una variable como dependiente y luego cambiandola a independiente, y sacando un promedio entre ambos resultados). Como resulta muy complejo obtener una ponderación entre ambos modelos (sea simétrica o no) queda propuesto.

Por último se tiene la comparación entre χ^2 . Se observa que para Gauss el valor es menor, lo que significa que la diferencia entre el modelo y los datos es mejor, pero no es nada concluyente ya que los valores son del orden de 10^{-35} . Por lo menos eso indica que 'leastsq' tiene un margen de tolerancia suficientemente pequeño, de manera que seguir optimizando entregaría poca diferencia en los parámetros que se obtengan.

4.2. La veracidad de los modelos

Dado que las figuras (2) y (3) no reflejan de forma suficientemente clara la superioridad de un modelo sobre otro, por lo que resulta importante realizar este test para ver que probabilidad es más probable (valga la redundancia) que haya generado los datos experimentales. La siguiente tabla muestra las diferencias máximas, críticas y la confianza. Notar que la diferencia crítica es redundante.

Cuadro 2: Diferencia máxima, crítica y nivel de confianza para cada modelo con respecto a los datos

	D _n	Nivel de Confianza	D _n crítico
Gauss	0.1647	0.0023	0.122
Lorentz	0.1656	0.0021	0.122

Comparando el D_n crítico y el obtenido para ambos modelos se deben rechazar ambos modelos. De la misma manera, el nivel de confianza refleja que es muy poco confiable. Al

observar las figuras (4) y (5) queda esto en evidencia: En varias zonas se aleja el modelo a los datos de forma considerable.

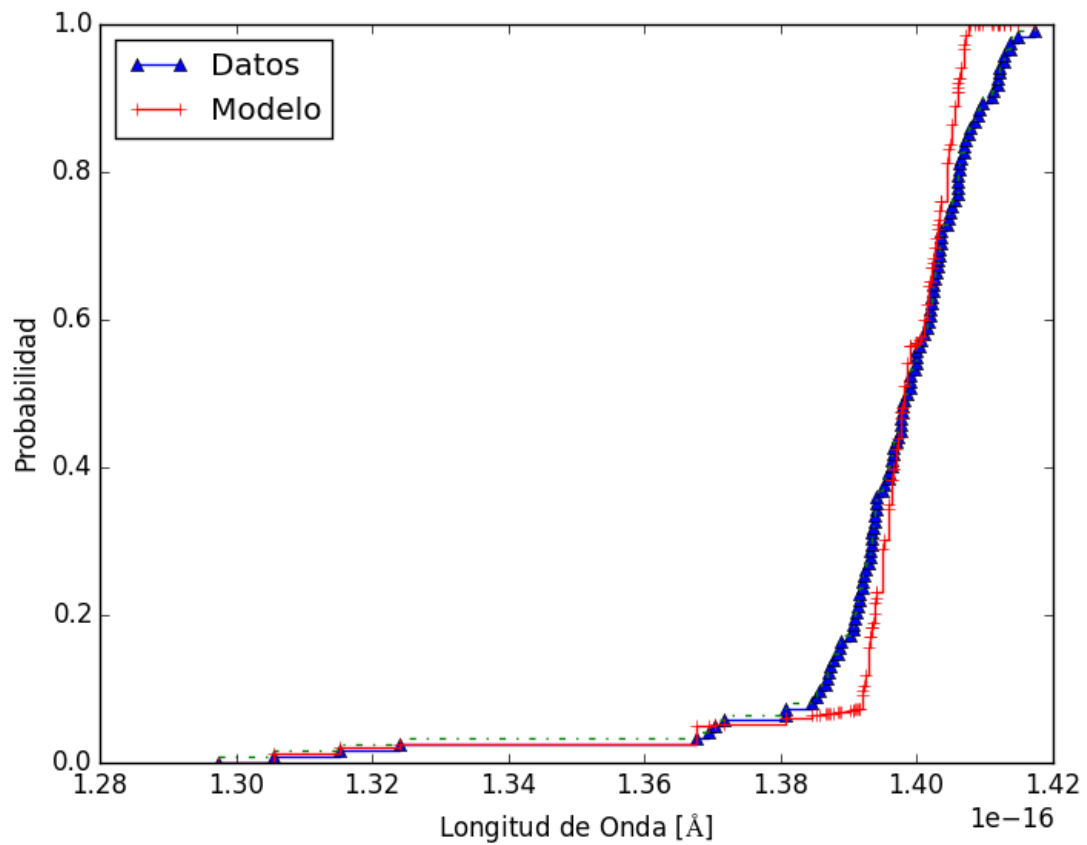


Figura 4: Función de densidad de probabilidad acumulada del modelo de Gauss (en rojo) y de los datos experimentales (en azul).

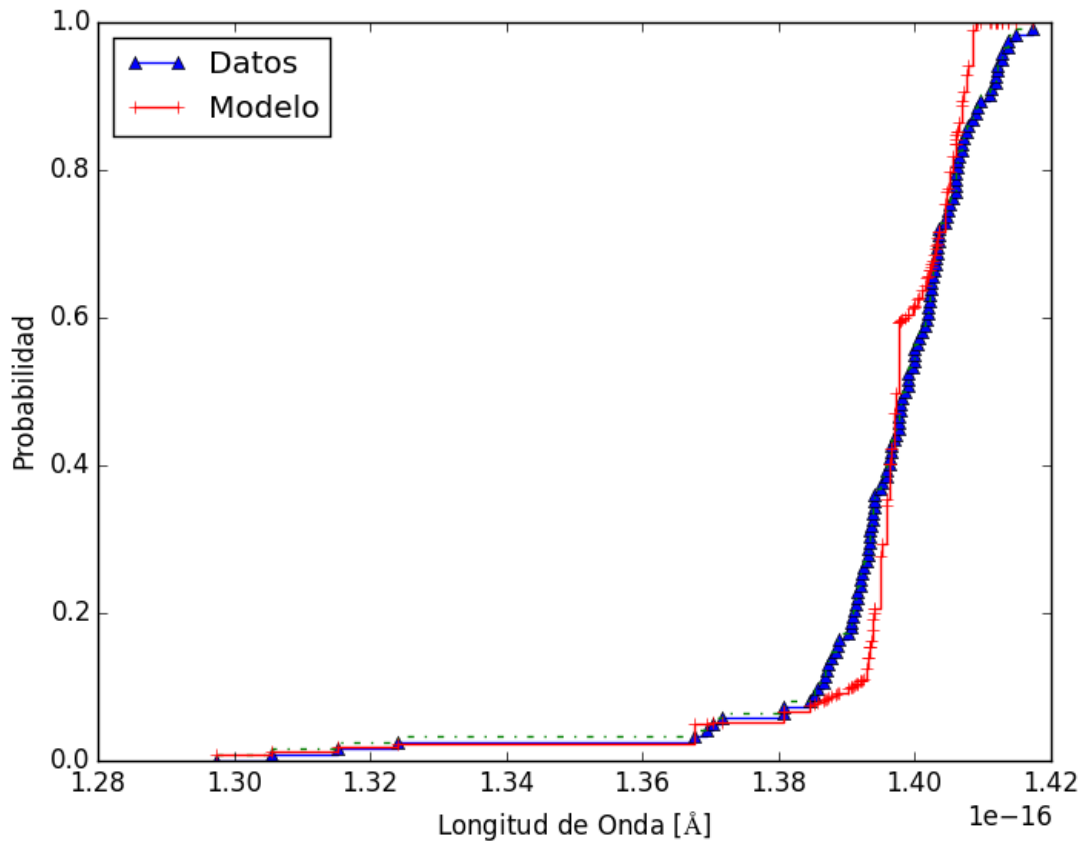


Figura 5: Función de densidad de probabilidad acumulada del modelo de Lorentz (en rojo) y de los datos experimentales (en azul).

A pesar de que ambos modelos son bastante malos, Podemos decir que el de Gauss es mejor según su nivel de confianza y que está más cerca del D_n crítico.

5. Conclusiones

Se concluye que, a pesar de verse razonables los modelos, un análisis más riguroso determina que los modelos no son suficientemente buenos. Se propone ver si una mezcla de ambos métodos logran una aproximación mejor. Se destaca la importancia de la prueba de Kolmogorovo-Smirnov porque permite salir del error al creer que un modelo puede estar bueno cuando no lo está, o al menos tener una noción de que si se va a asumir que un modelo es correcto para poder obtener datos más interesantes, considerar que se pueden generar errores debido a eso.

Se destaca que la modelación de parámetros en este caso era mucho más desafiante: eran modelos más complejos que lineales. Esto permite tener una noción más general del modelamiento de parámetros y como funciona. Se observó que si no se escoge un punto de partida suficientemente bueno para los parámetros, la aproximación fallará. De todas formas, el problema no era tan complejo como para perderse en otros mínimos que no sirvieran, por

ejemplo.

Se lograron mejoras para la definición de funciones de manera que sean óptimas, entendibles y reutilizables entre los programas. Queda propuesto mejorar el uso de strings para nombrar archivos, crear labels. Se realizó en una función, pero se pudo haber generalizado mucho más. Ha sido un problema en algunas ocasiones esto al punto de crear varios gráficos en una misma función para que queden guardados con nombres diferentes. Se espera que con algunas cosas aprendidas y aplicadas en este código se puedan aplicar a tareas siguientes (si es que las hay o:).

Referencias

- [1] Espectroscopia <https://es.wikipedia.org/wiki/Espectroscopia>
- [2] Informe 9 Metodos Numericos, Leonardo Leiva: 2.1. Algoritmo de Levenberg-Marquardt (LMA)
- [3] Demo Prueba de Kolmogorov-Smirnov <https://nbviewer.jupyter.org/urls/dl.dropboxusercontent.com/u/1036711/KS-test.ipynb>