



Universidade de Vigo



Universidad  
Carlos III de Madrid



POLITÉCNICA

Mestrado en Matemática Industrial

2018 – 2019

Traballo de Fin de Mestrado

# Parametrización de espectros en Resonancia Magnética Nuclear

Pablo Sánchez González

**Data presentación:** Luns, 4 de febreiro de 2019  
**Titor/a académico/a:** Andrés Prieto Aneiros  
**Empresa:** Mestrelab Research  
**Titor/a empresa:** Santiago López Ponte



Mestrelab Research  
chemistry software solutions



# Índice xeral

<b>Resumo</b>	<b>III</b>
<b>1. Introducción</b>	<b>1</b>
1.1. Descrición da empresa . . . . .	1
1.2. Motivación do traballo . . . . .	2
<b>2. Resonancia magnética nuclear</b>	<b>3</b>
2.1. Espín . . . . .	3
2.2. Magnetismo . . . . .	5
2.2.1. Efecto Zeeman . . . . .	6
2.2.2. Momento magnético . . . . .	6
2.2.3. Precesión do espín . . . . .	7
2.3. Sinal RMN . . . . .	9
2.3.1. Transformada de Fourier . . . . .	10
2.3.2. Espectro ideal . . . . .	15
2.4. RMN en dúas dimensións . . . . .	16
2.4.1. Sinal RMN . . . . .	17
2.4.2. Transformada de Fourier discreta bidimensional . . . . .	18
<b>3. Modelo e algoritmos</b>	<b>19</b>
3.1. Caso unidimensional . . . . .	19
3.1.1. Identificación dos picos . . . . .	20
3.1.2. Reconstrución do espectro . . . . .	25
3.2. Caso bidimensional . . . . .	26
3.2.1. Identificación dos picos . . . . .	27
3.2.2. Reconstrución do espectro . . . . .	30

<b>4. Resultados</b>	<b>35</b>
4.1. Caso unidimensional . . . . .	35
4.2. Caso bidimensional . . . . .	43
<b>5. Conclusiones</b>	<b>53</b>
<b>Bibliografía</b>	<b>55</b>

## Resumo

O presente texto é a síntese do traballo levado a cabo na empresa Mes-trelab Research no campo do procesamento de espectros en resonancia magnética nuclear (RMN) por parte do alumno Pablo Sánchez González. Trátase da implementación e ampliación do método descrito en [1] para a identificación de picos de radiación nos mencionados espectros. Dise ampliación, pois os algoritmos que en [1] se describen están limitados a espectros de dominio unidimensional, mentres que este traballo supón a xeneralización dos mesmos para que poidan servir para analizar espectros bidimensionais.

Ao longo destas páxinas farase unha breve descrición da física que hai detrás da RMN, xunto cunha detallada disertación acerca do método que se implementou. A linguaxe de programación que se empregou é *C++*, pola súa eficiencia e debido a que está orientada a obxectos. Para finalizar, expóranse os resultados acadados.



# Capítulo 1

## Introdución

### 1.1. Descrición da empresa

Mestrelab Research xorde como unha empresa de base tecnolóxica nacida do departamento de química orgánica da Universidade de Santiago de Compostela especializado no desenvolvemento de software de apoio no campo da química analítica e no da bioquímica. Dá soporte a investigación en laboratorios, empresas farmacéuticas, industria química e demais campos relacionados.

Os campos nos que se centra son os que seguen: procesado e análise de **Resonancia Magnética Nuclear (RMN)**, **Cromatografía de Líquidos e Gases** e **Espectrometría de Masas**; verificación e elucidación, tanto automática como manual, de datos espectroscópicos e espectrométricos; aplicacións metabonómicas; almacenamento, arquivado, xestión e minado de datos analíticos e aplicacións quimioinformáticas entre outros.

Unha pequena enumeración dos mercados que aborda: descubrimento e desenvolvemento de fármacos; metabonómica; desenvolvemento de materiais; industria petroquímica e alimenticia; química agrícola, do medio ambiente e sintética, tanto orgánica como inorgánica; química industrial e edulcorantes e aromatizantes.

Actualmente Mestrelab ten en nómina en torno a 40 empregados localizados en distintos países. Dispón de locais comerciais en China, Xapón, Francia, Corea, Polonia e India. A sede principal localízase en Santiago de Compostela. Para máis información acerca desta empresa pode consultarse a súa páxina web [www.mestrelab.com](http://www.mestrelab.com).

## 1.2. Motivación do traballo

No eido do tratamento de sinais en RMN hai xa moito escrito, pero os métodos que se soen empregar para a detección de picos de radiación e reconstrución de espectros, non son moi eficientes computacionalmente, xa que un espectro de resonancia magnética nuclear pode chegar a conter grandes volumes de datos. Un intento de mellorar os tempos de calculo e de procesado destes espectros ven descrito en [1], onde se detallan unha serie de algoritmos que, segundo os autores, veñen a mellorar as presentes prácticas neste campo. Isto ten un claro interese, tanto académico, pois describe novas técnicas e algoritmos que poden ser moi interesantes, como económico, pois é unha das principais liñas de negocio de Mestrelab Research.

Neste contexto, queda claro que unha correcta implementación de novos algoritmos que aforren tempo de cálculo á hora de detectar e parametrizar picos de radiación en RMN, pode resultar moi beneficiosa. É por iso que o cometido do presente traballo é, en primeiro lugar, tratar de poñer en práctica o descrito en [1], e, en segundo, ampliar, na medida do posible, as técnicas aí descritas para poder empregar o método en espectros de dúas dimensións, xa que o artigo restrínxese ao caso de espectros unidimensionais.

A nivel académico este proxecto resulta extremadamente interesante, sobre todo pola súa multidisciplinaria. Por unha banda, hai que comprender claramente toda a física que hai detrás da técnica da resonancia magnética nuclear, que, aínda que non sexa clave para abordar o problema, si será necesaria para comprender que se está a facer. Por outra, ao longo do proxecto será preciso adquirir un amplo coñecemento de diversas técnicas computacionais, tanto de algoritmos, como o de *Savitzky-Golay* (para máis información consúltese [2]), coma de linguaxes de programación, como *C++*.



## Capítulo 2

# Resonancia magnética nuclear

Neste capítulo descríbese a resonancia magnética nuclear ou RMN. Fara-se unha descrición da física que hai detrás desta técnica así como da técnica en si, para logo introducir brevemente a resonancia magnética nuclear bidimensional. Non se pretende nesta sección abordar a totalidade deste método, o cal non sería práctico pola súa extensión, se non dar a entender ao lector como funciona someramente e en que se basea en aras da completitude do presente traballo.

### 2.1. Espín

Para comprender como funciona este método débese entender que é o espín, é dicir, o momento angular intrínseco das partículas elementais. É preciso comprender como se comporta o momento angular a nivel cuántico, xa que, como se verá, pouco ten que ver coa concepción clásica que se soe ter do mesmo. Para máis información acerca deste tema pode consultarse [10].

Cando o paradigma co que se está a traballar é cuántico, o momento angular está cuantizado, é dicir, unicamente pode tomar certos valores discretos. Considérese, por exemplo, o caso dunha molécula diatómica que rota; o módulo do momento angular total toma un dos seguintes valores:

$$L_{tot} = [J(J + 1)]^{1/2}\hbar, \quad (2.1)$$

onde o número cuántico  $J$  é un enteiro,  $J = 1, 2, \dots$  e  $\hbar$  é a constante de Planck reducida que toma un valor de  $\hbar = 6,626 \cdot 10^{-34} J \cdot s$ .

A enerxía rotacional,  $E_J$ , dunha molécula é proporcional ao cadrado do momento angular total, neste caso

$$E_J = BJ(J + 1), \quad (2.2)$$

onde a constante de proporcionalidade  $B$  denomínase constante rotacional da molécula e toma valores pequenos para moléculas pesadas e grandes para moléculas lixeiras.

Coñecido o módulo do momento angular sábese como de rápido está a xirar a molécula, pero pouco ou nada se sabe acerca da dirección do xiro. Aquí a cuántica imponse un limiar ao noso coñecemento, impedíndo coñecer a totalidade da información referente a esta cuestión. A pesar disto, é posible coñecer unha das compoñentes do vector momento angular, a cal tamén toma valores discretos dados por

$$L_z = M_J \hbar. \quad (2.3)$$

$M_J$ , o número cuántico acimutal, ven a ser un dos seguintes números enteiros  $-J, -J + 1, \dots, J$ , podendo tomar  $2J + 1$  valores. A partir da ecuación (2.2) pódese afirmar que para calquera estado rotacional no que se atope o sistema para un mesmo  $J$ , dos  $2J + 1$  posibles, a enerxía é a mesma. Nese caso dise que os estados están dexenerados.

Unha vez analizado o momento angular dende un punto de vista cuántico non se comporta como cabería esperar nun sistema clásico, describírase o espín propiamente. O espín dunha partícula é un momento angular ou, polo menos, compórtase como tal. Pódese entender como o momento angular asociado ao xiro da partícula sobre si mesma, pero ao non ter as partículas fundamentais unha xeometría definida, trátase máis ben dunha magnitude fundamental, algo así como a masa ou a carga. Para o caso do espín empregaremos  $S$  no canto de  $J$  para designar ao número cuántico, así como  $M_S$  en vez de  $M_J$ .

Agora verase como se combina o momento angular de dous sistemas con números cuánticos  $J_1$  e  $J_2$  para obter o momento angular total. Este virá dado por  $[J_3(J_3 + 1)]^{1/2}$ , onde  $J_3$  toma un dos seguintes valores

$$J_3 = \begin{cases} |J_1 - J_2|, \\ |J_1 - J_2| + 1, \\ \vdots \\ |J_1 + J_2|. \end{cases} \quad (2.4)$$

Isto quere dicir que o momento angular total ten un número cuántico que pode ir dende a suma dos seus compoñentes ata a súa diferenza, en valor absoluto, podendo tomar tódolos valores enteiros intermedios.

Como caso particular, considérese un sistema con dúas partículas de espín  $1/2$  (Recordemos que o espín é un vector, xa que é un momento angular. Cando se di que unha partícula tes espín  $1/2$  quérese dicir que ese é o valor do número cuántico de espín). Existen dúas posibilidades para o espín total deste sistema. Ou ben  $S = 0$  ou  $S = 1$ . No primeiro caso dise que os espíns de ambas partículas canceláanse entre si e son «antiparalelos». Para  $S = 1$  ambos espíns superpóñense construtivamente e dise que son «paralelos». Polo xeral ambos estados teñen diferente enerxía, pero non se pode afirmar cal deles ten menos.

O estado  $S = 1$  denomínase usualmente *triplete* ( $S = 0$  é coñecido como *singlete*), pois como vemos está dexenerado dando lugar a tres estados se temos en conta o número cuántico acimutal  $M_S \in \{-1, 0, 1\}$ . Na seguinte sección analizarase como esta dexeneración se pode romper mediante a acción dun campo magnético, dando lugar ao que se coñece como efecto Zeeman.

No estudo de espíns nucleares o número cuántico denótase por  $I$ . A maior parte dos núcleos posúen espín. Tal é o caso do deuterio ou  ${}^2H$ , que contén un protón e un neutrón, cada un deles con espín  $1/2$ , sendo un caso idéntico ao descrito anteriormente. Deste xeito, tense unha configuración de *singlete* con  $I = 0$  e unha de *triplete*  $I = 1$ . Para o deuterio o *triplete* é o estado de menor enerxía ou fundamental. En xeral para un átomo arbitrario o estado fundamental pode ser calquera, aínda que existen certas normas nas que non se profundará por carecer de interese para o presente traballo.

## 2.2. Magnetismo

Neste traballo adoptarase unha descrición clásica dos campos electromagnéticos. Isto quere dicir que os mesmos virán completamente descritos por dous campos vectoriais definidos en todo punto do espazo, o campo eléctrico  $\vec{E}$  e o campo magnético  $\vec{B}$ . O comportamento de ambos campos é dictado polas ecuacións de Maxwell. Unha descrición máis detallada deste apartado pode atoparse en [10].

### 2.2.1. Efecto Zeeman

Os estados con número cuántico  $I$  distinto de 0, en xeral, están dexenerados. Cando se aplica un campo magnético externo os niveis de enerxía dexenerados desdóbranse en varias liñas estreitamente espazadas. Este efecto foi observado por vez primeira por Pieter Zeeman e débese a interacción entre o campo magnético e o momento magnético asociado ao espín. Na Figura 2.1

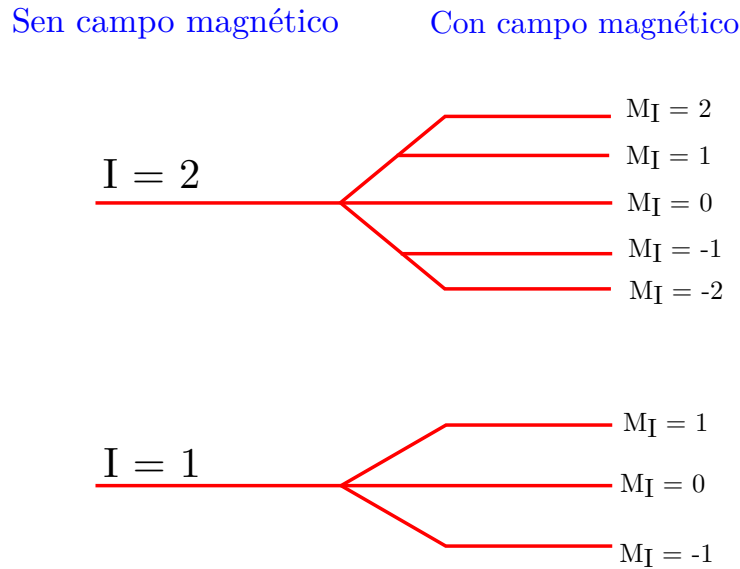


Figura 2.1: Desdobramento Zeeman. Nesta imaxe observase como os niveis enerxéticos dun núcleo se desdobran ante a presenza dun campo magnético.  $I$  é o número cuántico asociado ao espín nuclear, mentres que  $M_I$  é o número cuántico de espín

$M_I$  é o número cuántico acimutal asociado ao espín nuclear  $I$ .

### 2.2.2. Momento magnético

A magnitude que mide a interacción dunha sustancia cun campo magnético é o momento magnético  $\vec{\mu}$ , sendo a enerxía magnética dun sistema dependente da interacción desta cantidade co campo  $\vec{B}$ , definida por

$$E_{mag} = -\vec{\mu} \cdot \vec{B}. \quad (2.5)$$

A partir desta expresión pódese comprobar que a enerxía magnética é menor se  $\vec{\mu}$  e  $\vec{B}$  teñen a mesma dirección. Para a obter a enerxía magnética dun sistema débese integrar a ecuación (2.5) sobre o volume do mesmo.

A nivel cuántico, o espín está relacionado co momento magnético dunha partícula. Do teorema de Wigner-Eckart pódese deducir que o espín e o momento angular son directamente proporcionais a través dun coeficiente  $\gamma$  denominado factor xirromagnético, isto é

$$\tilde{\mu} = \gamma \tilde{S}. \quad (2.6)$$

$\tilde{\mu}$  e  $\tilde{S}$  son operadores cuánticos asociados ao momento magnético e ao espín, respectivamente. O factor xirromagnético ten signo e para partículas con  $\gamma$  positivo, o momento magnético ten o mesmo sentido que o espín, mentres que cando  $\gamma$  toma un valor negativo, teñen sentidos opostos.

### 2.2.3. Precesión do espín

O espín dunha partícula é un vector, como todo momento angular, que pode apuntar en calquera dirección do espazo. Tendo en conta (2.6), o momento magnético dun núcleo apunta, ou ben no mesmo sentido que o espín (se  $\gamma > 0$ ), ou ben no contrario (se  $\gamma < 0$ ). Tomando unha mostra en equilibrio, a distribución de momentos magnéticos será isotrópica, de maneira que non hai un momento magnético neto.

Agora ben, se sobre esta mostra é aplicado un campo magnético os espíns das partículas comezaran un movemento de precesión arredor do campo, mantendo un ángulo constante co mesmo. Isto débese a que teñen momento angular e magnético ao mesmo tempo.

#### 2.2.3.1. Frecuencia de Larmor

Unha análise cuántica da precesión do espín mostra que a frecuencia de precesión é

$$\omega^0 = -\gamma B^0, \quad (2.7)$$

onde  $B^0$  é o campo magnético na posición da partícula e  $\gamma$  é o factor xirromagnético, que xa foi introducido en (2.6). Denomínase a  $\omega^0$  frecuencia de Larmor. Esta frecuencia ten signo ben definido, o cal nos di o sentido de xiro do espín.

### 2.2.3.2. Paramagnetismo nuclear

No caso particular dunha mostra de  $^1H$  disolta en auga, se se acende, de xeito súbito, un campo magnético homoxéneo e intenso, os espíns dos núcleos de  $^1H$  comezarán a precesar arredor deste, pero dado que a orientación dos mesmos é isotrópica, non aparecerá un momento magnético neto, polo menos non de xeito instantáneo. Para unha descrición máis detallada deste comportamento consultar [10].

Sucedede que, dado que os protóns da mostra están rodeados de moléculas de auga, existe certa interacción magnética dos mesmos co medio. Isto fai que o eixo en torno ao que precesa cada protón varíe lixeiramente co tempo chegando a producirse un equilibrio térmico no cal se acada un estado de mínima enerxía. Isto rompe a isotropía do medio, pois os estados nos que o momento magnético nuclear e o campo magnético externo están aliñados son máis estables. Desta maneira chega a producirse un momento magnético neto na mostra.

### 2.2.3.3. Magnetización transversal

Coñécese como magnetización  $\vec{M}$ , á densidade de volume de momento magnético e é a unidade para dar conta desta magnitude a escala macroscópica. Matematicamente pódese expresar como a derivada do momento magnético con respecto ao volume:

$$\vec{M} = \frac{d\vec{\mu}}{dV}. \quad (2.8)$$

Acábase de describir que ao cabo dun tempo tras aplicarlle un campo magnético á mostra, aparece un momento magnético neto lonxitudinal ao campo aplicado. Na práctica a magnetización asociada a el é case indetectable. O que se fai normalmente é empregar un pulso de radiofrecuencias que xira tódolos espíns da mostra nun ángulo de  $\pi/2$  arredor dun eixo determinado. Desta maneira, a polarización neta dos espíns apuntará agora nunha dirección transversal ao campo magnético aplicado.

Desta maneira, agora a magnetización neta da mostra apuntará nunha dirección ortogonal ao campo magnético aplicado. O que sucede neste momento é que se apaga o pulso electromagnético que xirou a polarización da mostra. Desta maneira os espíns comezarán de novo a precesar en torno ao campo aplicado e, ao xirar estes, tamén o fai o momento magnético neto e a magnetización coa mesma frecuencia, é dicir, a de Larmor. Lentamente a

mostra retorna ao seu estado de equilibrio, pero esta magnetización transversal pode ser detectada e medida dado que oscila a unha frecuencia moi ben definida.

## 2.3. Sinal RMN

Ocorre que un momento magnético que rota xera un campo magnético variable no tempo. Situando unha bobina nas súas proximidades, este creará un campo eléctrico que comezará a mover os electróns dando lugar a unha corrente a través do cable. Esta corrente é o que se detecta dando lugar a un sinal RMN ao cal denominaremos FID polas súas siglas en inglés (*free-induction decay*).

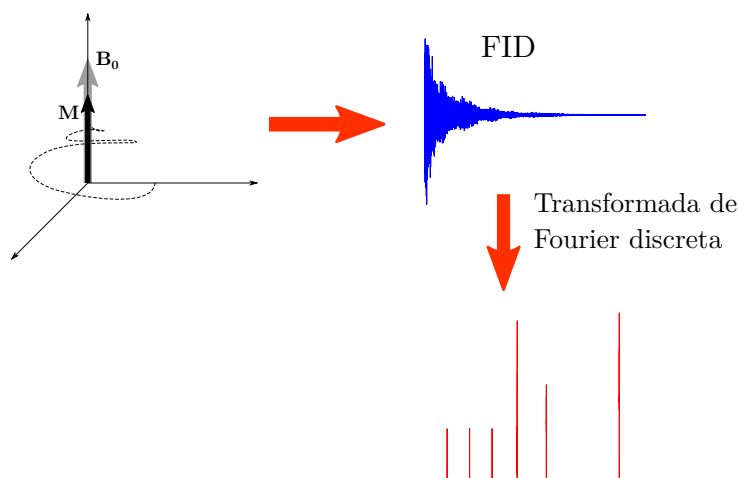


Figura 2.2: Arriba a esquerda pode verse unha representación do movemento da magnetización  $\mathbf{M}$  en torno ao campo magnético  $\mathbf{B}_0$ . A continuación mostrase o *FID* (free induction decay) ou sinal medido. Por último, tras aplicarlle a transformada de Fourier discreta a este sinal, represéntase o seu espectro. Imaxe inspirada na Figura 2.5 de [13]

Na Figura 2.2 pode observarse á esquerda a representación da magnetización  $\bar{M}$  da mostra oscilando en torno ao campo inducido  $B_0$ . De seguido está a FID, é dicir, o sinal eléctrico co que se traballará e, por último, a transformada de Fourier deste sinal, que non é máis que unha representa-

ción do mesmo no espazo de frecuencias no canto do temporal. De seguido explicarase con claridade en que consiste a transformada de Fourier.

### 2.3.1. Transformada de Fourier

A transformada de Fourier é unha transformación matemática reversible empregada para transformar sinais entre o dominio temporal e o dominio das frecuencias. Para un estudo en profundidade desta operación matemática no contexto da RMN é recomendable consultar [8]. A transformada de Fourier dunha función integrable  $f : \mathbb{R} \rightarrow \mathbb{C}$  ven dada por

$$\hat{f}(\xi) = \int_{-\infty}^{\infty} f(t)e^{-2\pi it\xi} dt. \quad (2.9)$$

No caso de que a variable independente  $t$  representase o tempo, a variable transformada  $\xi$  representaría a frecuencia. A transformada inversa permite recuperar  $f$  a partir de  $\hat{f}$

$$f(t) = \int_{-\infty}^{\infty} \hat{f}(\xi)e^{2\pi it\xi} d\xi, \quad (2.10)$$

para calquera número real  $t$ .

Cando se adquire un sinal RMN rexístrase no dominio temporal, que é o que ten un sentido físico á hora de medir, pero para traballar con el é máis útil pasar ao dominio das frecuencias, para o cal se emprega a transformada de Fourier (véxase a Figura 2.2). Sempre que se fale dun espectro RMN estarase a representar a FID no dominio das frecuencias.

#### 2.3.1.1. Transformada de Fourier discreta

Acábase de ver unha definición da transformada de Fourier que non se pode aplicar para o procesamento de sinais en RMN. Isto débese a que as variables que aparecen nas ecuacións anteriores (2.9) e (2.10), son continuas, mentres que os datos reais son discretos, posto que tódolos instrumentos modernos rexístran o sinal RMN en intervalos temporais discretos, almacenando os datos en series temporais que se poden representar como un vector,

$$\vec{f} = (f_0, f_1, \dots, f_{N-1}), \quad (2.11)$$

sendo  $N$  o número total de medidas e onde os subíndices indican a orde temporal da medida. Polo xeral o intervalo entre medidas é constante, é



dicir,  $f_i = f(t_i)$ , onde  $t_i = t_0 + i\delta$  para algunha constante  $\delta$  que marca o espazado entre os datos e  $i = 0, 1, 2, \dots, N - 1$ .

Así pois, ao ser o sinal temporal un conxunto de datos discretos, tamén o será o espectro de frecuencias, o cal representaremos por

$$\vec{\hat{f}} = (\hat{f}_0, \hat{f}_1, \dots, \hat{f}_{N-1}), \quad (2.12)$$

onde agora  $N$  é o número de compoñentes do espectro. Defínese, pois, a transformada de Fourier discreta, que será a que permite pasar da serie temporal discreta ao espectro discreto de frecuencias, como

$$\hat{f}_n = \frac{1}{\sqrt{N}} \sum_{k=0}^{N-1} f_k e^{-2\pi i k n / N} \quad \text{para } n = 0, \dots, N - 1. \quad (2.13)$$

Débese notar que, en ocasións, esta transformada defínese de xeitos diferentes, sexa con outras constantes de normalización, sexa con outro signo no argumento, pero estas variantes non afectan as propiedades desta operación (para máis información consúltase [8]).

Do mesmo xeito que antes, no caso discreto tamén se pode definir a transformada inversa, que nos devolverá ao dominio temporal se estamos no das frecuencias, mediante a relación

$$f_k = \frac{1}{\sqrt{N}} \sum_{n=0}^{N-1} \hat{f}_n e^{2\pi i k n / N} \quad \text{para } k = 0, \dots, N - 1. \quad (2.14)$$

### 2.3.1.2. Teorema da convolución

Agora describirase un importante teorema que permitirá entender a forma que toman os espectros RMN e porqué son modelados os picos de radiación como funcións de Lorentz. O teorema establece que a transformada de Fourier discreta do produto elemento a elemento (denotado por  $\odot$ ) de dous vectores  $\vec{f}$  e  $\vec{a}$  ven dada pola convolución da transformada de  $\vec{f}$  coa transformada de  $\vec{a}$ , isto é,

$$[\text{DFT}(\vec{f} \odot \vec{a})]_n = \frac{1}{\sqrt{N}} \sum_{j=0}^{N-1} \hat{f}_j \hat{a}_{n-j} \quad \text{para } n = 0, \dots, N - 1, \quad (2.15)$$

onde  $\vec{\tilde{a}}$  é a transformada de Fourier discreta de  $\vec{a}$ . Analogamente, aplicando a transformada inversa sobre o produto compoñente a compoñente das transformadas, obtemos a convolución das series temporais

$$[\text{IDFT}(\vec{\tilde{f}} \odot \vec{\tilde{a}})]_k = \frac{1}{\sqrt{N}} \sum_{j=0}^{N-1} f_j a_{k-j} \quad \text{para } k = 0, \dots, N-1. \quad (2.16)$$

Nótese que os produtos denotados por  $\odot$  entre vectores das ecuacións (2.15) e (2.16) non son produtos escalares. O que se está a facer é multiplicar os elementos de ambos vectores que ocupan a mesma posición para obter un terceiro (esta operación tamén é coñecida como produto de Hadamard entre vectores).

### 2.3.1.3. Transformadas de Fourier notables

Falaremos agora da transformada de dúas funcións en concreto que serven para modelizar un espectro RMN ideal.

**Delta de Kronecker  $\delta_m$  centrada no punto  $x_m$  con  $0 \leq m \leq N-1$ :** As compoñentes do vector que determinan os valores discretos veñen dados polo valor cero a excepción da compoñente  $m$ -ésima, que toma o valor un. Se o punto que toma o valor unidade é o primeiro da ecuación (2.13) dedúcese que a transformada discreta de Fourier é a función constante con valor  $1/\sqrt{N}$  en todo o dominio. Se o valor non nulo se atopa noutro lugar, a súa transformada será unha función sinusoidal con frecuencia  $2\pi/m$  sendo  $m$  o índice do punto que toma o valor non nulo. Na Figura 2.3 pódese ver unha representación gráfica disto.

**Decaemento exponencial:** Teoricamente, un sinal FID consiste en sinusoides que decaen exponencialmente co tempo. Considérese o caso máis simple, un senoide con frecuencia cero, é dicir, unha función  $f(t)$  que é nula para  $t < 0$  e que ven dada por unha exponencial decrecente cun ratio de  $\pi L$ . A súa tranformada de Fourier continua ven dada por

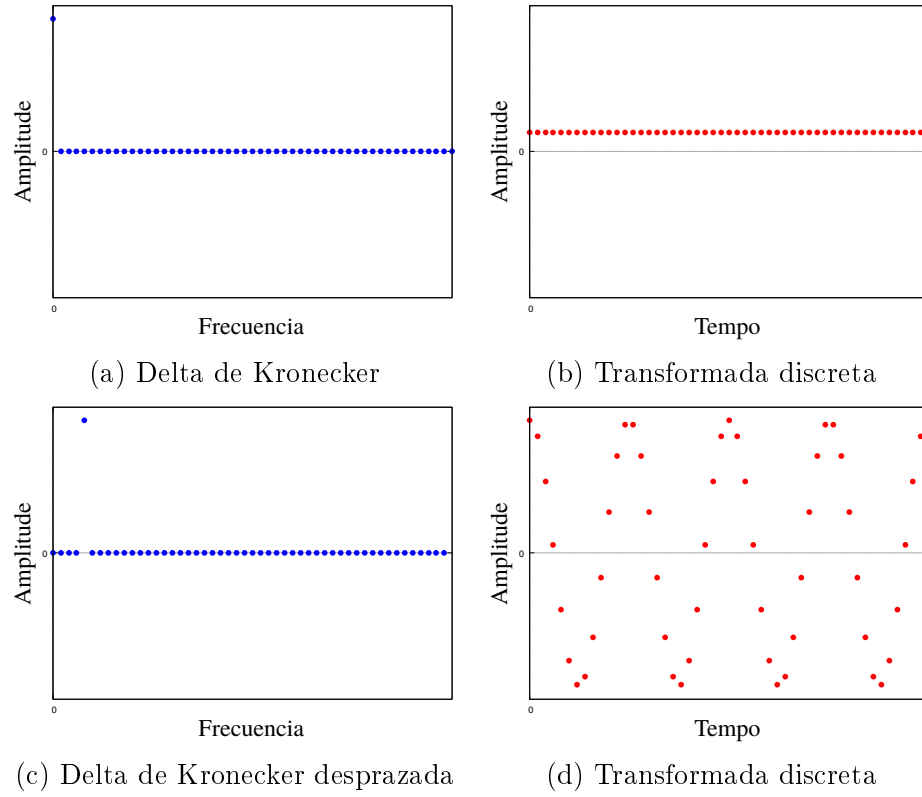


Figura 2.3: Transformada discreta da Delta de Kronecker. Arriba a esquerda vese a Delta centrada no cero. A dereita desta móstrase a súa transformada discreta. Abaixo a esquerda pode verse.

$$\begin{aligned}
 \hat{f}(\omega) &= \int_0^\infty e^{-\pi L t} e^{-2\pi i \omega t} dt = \int_0^\infty e^{-(\pi L + 2\pi i \omega)t} dt \\
 &= \frac{1}{\pi L + 2\pi i \omega} \\
 &= \frac{\pi L}{(\pi L)^2 + (2\pi \omega)^2} - i \frac{2\pi \omega}{(\pi L)^2 + (2\pi \omega)^2}. \quad (2.17)
 \end{aligned}$$

A parte real toma como valor máximo  $1/\pi L$ , o cal ocorre en  $\omega = 0$ . A metade deste valor acádase cando  $\omega = \pm L/2$ , razón pola cal  $L$  se denomina FWHM (full width at half maximum) polas súas siglas en inglés, ou ancho total á metade do máximo, ao longo deste traballo tamén se introducirá a

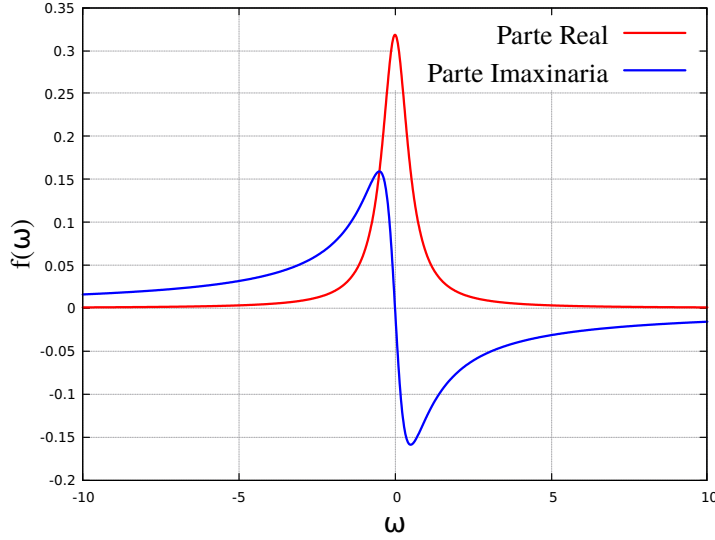


Figura 2.4: Compoñentes da curva de Lorentz. En azul pode verse a parte imaxinaria mentres que en vermello a parte real da ecuación (2.17).

cantidade  $\lambda$ , que é a metade de anchura á metade de altura. A expresión (2.17) é coñecida como *Lorentziana*. Para máis información pódese consultar [8].

A parte real, que é a que se soe representar, denomínase parte absorbente. Por outra banda, a parte imaxinaria dise parte dispersiva. Ámbalas dúas son representadas na Figura 2.4. Un dos motivos polos que usualmente se representa a parte real é porque ao ser estreita e simétrica, nun espectro moi poboado, é menos probable que haxa solapes entre distintos picos. Ademais, como se verá despois, é posible eliminar a parte imaxinaria.

Por último, móstrase como un vector de valores discretos con decaemento exponencial é transformado mediante a transformada de Fourier discreta. Neste caso,  $f_k = f(t_k) = e^{-\pi Lk\Delta t}$  e resulta

$$\hat{f}_n = \frac{1}{\sqrt{N}} \sum_{k=0}^{N-1} e^{-\pi Lk\Delta t} e^{-2\pi i kn/N} = \frac{1}{\sqrt{N}} \sum_{k=0}^{N-1} z^k, \quad (2.18)$$

onde  $z = e^{-\pi L\Delta t - 2\pi i n/N}$ , e  $\Delta t$  é a distancia de equiespaciado entre os puntos onde se avalía a función con decaemento exponencial. Así pois, estamos ante unha serie xeométrica, co cal

$$\hat{f}_n = \frac{1}{\sqrt{N}} \frac{1 - z^N}{1 - z} = \frac{1}{\sqrt{N}} \frac{1 - e^{-\pi LN\Delta t - 2\pi in}}{1 - e^{-\pi L\Delta t - 2\pi in/N}}. \quad (2.19)$$

Cando  $\pi LN\Delta t$  é grande, é dicir, cando o decaemento da exponencial é rápido comparado co valor de  $t_{N-1}$ , tense que

$$\hat{f}_n = \frac{1}{\sqrt{N}} \frac{1}{1 - e^{-\pi L\Delta t - 2\pi in/N}}. \quad (2.20)$$

### 2.3.2. Espectro ideal

Considérese agora unha función sinusoidal que decae exponencialmente de frecuencia distinta de cero  $\omega_0$ :

$$f(t) = e^{-\pi Lt} e^{2\pi i\omega_0 t}. \quad (2.21)$$

A transformada de Fourier continua resulta ser unha curva de Lorentz centrada en  $\omega_0$ , que seguindo cálculos análogos aos feitos en (2.16), obtense

$$\hat{f}(\omega) = \frac{\pi L}{(\pi L)^2 + (2\pi(\omega - \omega_0))^2} - i \frac{2\pi(\omega - \omega_0)}{(\pi L)^2 + (2\pi(\omega - \omega_0))^2}. \quad (2.22)$$

O teorema da convolución da transformada de Fourier continua (véxase [8]) permite entender o espectro dunha función sinusoidal que decae exponencialmente como a convolución dunha delta de Dirac centrada en  $\omega_0$  cunha Lorentziana de anchura  $L$  (aínda que este resultado excede o material descrito en este traballo ao tratarse dunha convolución dunha distribución cunha función integrable).

Destá maneira, é posible modelar un espectro de xeito ideal como unha superposición (combinación lineal) de curvas Lorentzianas, cada unha co seu centro, amplitude e anchura correspondente, isto é, dada un sinal FID  $s(t)$  (con valores nulos para  $t < 0$ ), a súa transformada de Fourier continua  $\hat{s}(\omega)$  pode ser modelada como a suma de  $J$  curvas Lorentzianas.

$$S(\omega) = \int_0^\infty s(t) e^{-i\omega t} dt = \sum_{j=0}^J e^{i\varphi_j} (a_j(\omega) + id_j(\omega)), \quad (2.23)$$

onde  $S(\omega)$  representa o espectro (empregarase  $S(\omega)$  para representar os espectros no canto de  $\hat{s}$ ). De feito, farase referencia a eles continuamente non

simplemente en tanto a súa condición de transformada de Fourier dun sinal FID. O coeficiente

$$a_j(\omega) = A_j \frac{T_{2j}^*}{1 + (\omega - \omega_j)^2 T_{2j}^{*2}} \quad (2.24)$$

é a compoñente absorbente e

$$d_j(\omega) = A_j \frac{T_{2j}^*(\omega - \omega_j)}{1 + (\omega - \omega_j)^2 T_{2j}^{*2}} \quad (2.25)$$

é a parte dispersiva. Nas expresións (2.24) e (2.25),  $\omega_j$  é a frecuencia do centro da curva Lorentziana,  $T_{2j}$  é o ratio de decaemento,  $A_j$  é a amplitude da mesma e  $\varphi_j$  é a fase espectral da compoñente  $j$ . O superíndice  $*$  é o operador de conxugación de valores complexos.

Da ecuación (2.23) deducimos que cunha corrección da fase  $\varphi_j$  adecuada, poderemos eliminar a parte imaxinaria do análise. Isto é o que se fai na práctica. Desta maneira que se poderá modelar un espectro ideal do seguinte xeito:

$$S(\omega) = \sum_{j=0}^J Y_j(\omega), \quad (2.26)$$

con

$$Y_j(\omega) = A_j \frac{\lambda_j}{\lambda_j^2 + (\omega - \omega_j)^2}, \quad (2.27)$$

onde  $\lambda_j = 1/T_{2j}^*$  é o parámetro HWHH (metade de anchura á metade de altura) xa introducido no estudo da transformada de Fourier dunha función con decaemento exponencial.

## 2.4. RMN en dúas dimensións

Dado que o propósito deste traballo non é profundar nos aspectos teóricos da resonancia magnética nuclear, senón desenvolver un programa que, dado un espectro, sexa quen de reconstruílo en base ao modelo que se describirá no seguinte capítulo, e debido a extensión que podería acadar este apartado de facer a descrición en detalle, farase simplemente unha breve introdución desta técnica aplicada a dominios bidimensionais. Pode atoparse máis información sobre a resonancia magnética nuclear en dúas dimensións en [3] e [11].

### 2.4.1. Sinal RMN

Un espectro bidimensional obtense ao aplicarlle a transformada de Fourier a un sinal  $s(t_1, t_2)$  que depende de dúas variables temporais  $t_1$  e  $t_2$ . Isto sucede debido a que o eixo temporal se divide en tres anacos, sendo  $t_1$  a variable que mide o tempo do segundo intervalo temporal e  $t_2$  a que fai o propio no terceiro intervalo. Na Figura 2.5 mostrase esta división temporal.

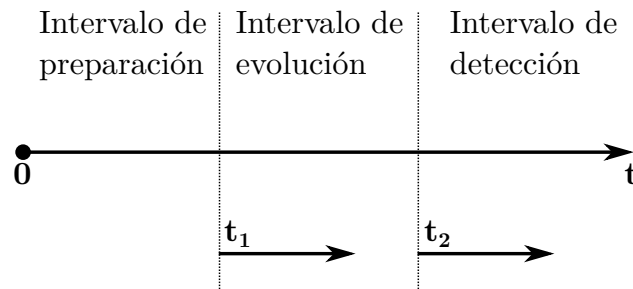


Figura 2.5: División do eixo temporal en RMN bidimensional. A variable  $t_1$  da conta da duración do intervalo de evolución, mentres que  $t_2$  é a variable que mide o tempo no intervalo de detección.

**Intervalo de preparación:** Nun primeiro momento o sistema é preparado para estar nun estado inicial adecuado. Igual que se facía no caso unidimensional, a mostra sitúase baixo a acción dun campo magnético ata que aparece unha magnetización neta, para posteriormente xirar a mesma mediante un pulso de radiofrecuencias.

**Intervalo de evolución:** Posteriormente, e a diferenza do caso unidimensional, deixase evolucionar a magnetización resultante durante un tempo  $t_1$ . Esta é a primeira das variables temporais do espectro. Durante este proceso non se rexistra o sinal aínda. Para contar con varios valores da variable  $t_1$  haberá que repetir o proceso tantas veces como sexa necesario con distintas duracións do intervalo de evolución.

**Intervalo de detección:** Por último, aplícaselle un novo pulso á mostra. É entón cando se mide o sinal ou *free induction decay* do mesmo xeito que se facía no caso unidimensional, sendo  $t_2$  a variable temporal.

Cabe destacar que a dimensión temporal correspondente a  $t_1$  se coñece como indirecta e, debido a que para cada valor temporal rexistrado na mesma se require repetir completamente o proceso con outro período de evolución, ten unha resolución moito menor a  $t_2$ , pois a cantidade de valores que tomará  $t_1$  é bastante máis pequena. A dimensión que corresponde a  $t_2$  coñécese, loxicamente, como directa.

### 2.4.2. Transformada de Fourier discreta bidimensional

Nesta sección descríbese moi brevemente a transformada de Fourier discreta no caso no que o dominio teña dúas dimensións e forma rectangular. Simplemente será introducida esta transformada discreta xunto coa súa inversa. Neste caso, á diferenza do caso unidimensional, os datos temporais serán almacenados nunha matriz de dúas dimensións  $M \times N$ , de maneira que  $[\mathbf{F}]_{hk} = f_{hk}$  almacene o valor  $k$ -ésimo que toma  $t_2$  para o valor  $h$ -ésimo de  $t_1$ .

Así, a partir dun sinal temporal bidimensional, pode obterse o espectro de frecuencias aplicando a transformada de Fourier discreta en dúas dimensións, denotada por  $\hat{\mathbf{F}}$ , que ven definida por

$$[\hat{\mathbf{F}}]_{mn} = \hat{f}_{mn} = \frac{1}{\sqrt{MN}} \sum_{h=0}^{M-1} \sum_{k=0}^{N-1} f_{hk} e^{-2\pi i(kn/N + hm/M)}, \quad (2.28)$$

para  $0 \leq m \leq M - 1$ ,  $0 \leq n \leq N - 1$ . No caso de ter un espectro e querer recuperar o sinal bidimensional pode aplicarse a transformada inversa que se mostra na ecuación (2.29).

$$[\mathbf{F}]_{hk} = f_{hk} = \frac{1}{\sqrt{MN}} \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} \hat{f}_{mn} e^{2\pi i(kn/N + hm/M)}, \quad (2.29)$$

para  $0 \leq h \leq M - 1$ ,  $0 \leq k \leq N - 1$ .



# Capítulo 3

## Modelo e algoritmos

Neste capítulo introducírase o algoritmo que se describe en [1] e que ven a vertebrar este traballo. En resumidas contas, trátase de identificar picos de radiación no dominio das frecuencias para un espectro RMN para posteriormente tratar de describilos de xeito analítico mediante funcións Lorentzianas. En [1] defínese o método numérico para o caso dun espectro unidimensional, pero o cometido principal deste traballo de fin de mestrado é a xeneralización do método para que funcione con espectros bidimensionais. Nas seguintes seccións explícase tanto o caso unidimensional como o bidimensional.

### 3.1. Caso unidimensional

Pódese dividir o método en dúas partes ben diferenciadas, por un lado a identificación dos picos de radiación, e por outro a reconstrución do espectro mediante funcións analíticas.

Por identificar os picos de radiación entenderase o proceso para localizalos no espectro, é dicir, atopar a frecuencia á que se producen. Para isto emprégase a segunda derivada do espectro. Localizando un mínimo local na segunda derivada, estase ante unha curvatura máxima (na parte cóncava da curva) por parte do espectro, que é o que interesa buscar para encontrar os picos. Na Figura 3.1 mostrase un esquema desta situación.

Unha vez feito isto e, dado que, por norma xeral, os espectros están contaminados con ruído, débense filtrar os picos potenciais que se acaban de identificar. Para iso o artigo describe unha función que puntuía cada posible pico. Despois de puntuados so resta establecer un limiar de puntuación por

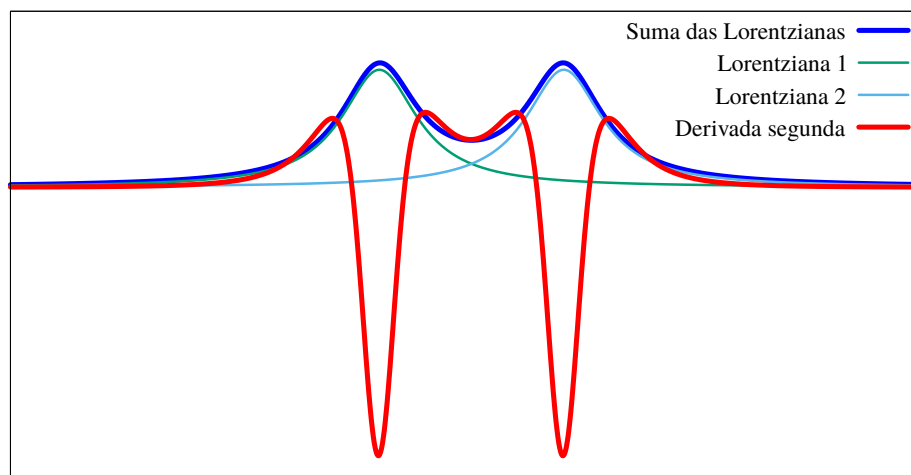


Figura 3.1: Mínimos locais da segunda derivada. En azul móstrase a suma das curvas *Lorentziana 1* e *Lorentziana 2*. En vermello, a derivada segunda desta suma.

baixo do cal se considerará que eses hipotéticos picos son debidos ao ruído.

Chegados a este punto os picos están debidamente identificados e localizados. O que resta agora é buscar a mellor aproximación de cada un deles a unha Lorentziana, para desa maneira, reconstruír de xeito aproximado o espectro experimental, ou sintético, como suma desas funcións analíticas. Para iso séguese un proceso iterativo no cal se van aproximando as correspondentes curvas Lorentzianas asociadas a cada pico ata que describan o espectro o máis fielmente posible.

### 3.1.1. Identificación dos picos

No que segue detállase con máis claridade como se procede para a identificación dos picos da transformada de Fourier dun sinal FID. Primeiramente débese calcular a segunda derivada do espectro pois, como xa se viu, será a que indique a posición dos mesmos. Para o cálculo da segunda derivada empregouse un método diferente ao empregado en [1], pois neste a derivada é aproximada simplemente mediante diferencias finitas, o cal non é adecuado para sinais con ruído. Para corrixir o efecto do ruído, posteriormente aplican un filtro de media, o cal promedia os valores próximos para atenuar as fluctuacións estocásticas debidas ao ruído. Existe un método máis sinxelo e potente computacionalmente que consiste na a aplicación dun filtro

de Savitzky-Golay (para profundar neste tema consultar [2] e [7]), o cal se describe ao final desta sección.

Unha vez calculada numericamente a derivada segunda do espectro, xérase unha lista con tódolos mínimos locais, xa que estes serán os que nos marquen a posición dos picos do espectro. O problema está en que o sinal está contaminado con ruído e, aínda empregando o filtro de Savitzky-Golay, débense filtrar os picos de radiación das fluctuacións arbitrarias debidas a outros factores. Con este fin, en [1] descríbese unha función que puntuá cada un dos mínimos locais da segunda derivada do espectro, para logo seleccionar só aqueles que teñan unha puntuación superior a un certo valor.

O algoritmo proposto consiste en asociarlle a cada mínimo local un pequeno intervalo do dominio, definindo deste xeito unha tripla de puntos (os extremos do intervalo, xunto co propio mínimo, que debe quedar no interior do mesmo). Para definir este intervalo tomaranse por cada extremo ou ben o punto de corte co eixo horizontal máis próximo, ou ben o máximo local máis próximo (destes dous, aquel que sexa máis achegado ao mínimo) da segunda derivada. Así, se  $\omega_m$  é o mínimo, constrúese a tripla  $(\omega_l, \omega_m, \omega_r)$  do seguinte xeito:

$$\begin{aligned}
 & \omega_m \text{ é tal que } S''(\omega_m) < 0 \text{ e } S''(\omega_{m-1}) > S''(\omega_m) < S''(\omega_{m+1}), \\
 & \omega_l = \operatorname{argmin}_{0 \leq j < m} \left( \omega_j - \omega_m \left| \left( \overbrace{S''(\omega_{j-1}) \geq 0 \text{ e } S''(\omega_j) < 0}^{\text{Raíz máis próxima}} \right. \right. \right. \\
 & \qquad \qquad \qquad \left. \left. \left. \text{ou } \underbrace{S''(\omega_{j-1}) \leq S''(\omega_j) > S''(\omega_{j+1})}_{\text{Máximo local máis preto}} \right) \right) \right) \\
 & \omega_r = \operatorname{argmin}_{m < j < N-1} \left( \omega_m - \omega_j \left| \left( \overbrace{S''(\omega_{j+1}) \geq 0 \text{ e } S''(\omega_j) < 0}^{\text{Raíz máis próxima}} \right. \right. \right. \\
 & \qquad \qquad \qquad \left. \left. \left. \text{ou } \underbrace{S''(\omega_{j-1}) < S''(\omega_j) \geq S''(\omega_{j+1})}_{\text{Máximo local máis preto}} \right) \right) \right) \quad (3.1)
 \end{aligned}$$

Nótese que os puntos do dominio estanse a denotar coa variable  $\omega$  (téñase en conta que as restriccións que involucran os valores  $\omega_{-1}$  e  $\omega_N$  non se terán en conta ao atoparse fora do dominio discreto considerado).  $S''(\omega)$  é a derivada

segunda do espectro. Máis adiante, na sección 3.1.1.1, explícase en detalle como calculala na práctica. Tense, pois, un conxunto  $\Omega$  formado por todas as triplas  $\mathbf{p} = (\omega_l, \omega_m, \omega_r)$ , sendo  $\omega_m$  o punto do dominio no que se localiza o mínimo e  $\omega_l$  e  $\omega_r$  os extremos esquerdo e dereito do intervalo, respectivamente. A continuación procederase a puntuar cada unha delas mediante a seguinte función:

$$\text{puntuación}(\mathbf{p}) = \frac{\min \left( \sum_{k=l}^m |S''(\omega_k)|, \sum_{k=m}^r |S''(\omega_k)| \right)}{\max_{\mathbf{q} \in \Omega} \text{puntuación}(\mathbf{q})}. \quad (3.2)$$

Agora que as triplas están debidamente clasificadas con esa puntuación, débese establecer a puntuación mínima para ser consideradas picos de radiación. En [1] emprégase para tal propósito a puntuación media máis un parámetro  $\delta$  multiplicado pola desviación típica das puntuacións nunha zona do espectro na que se saiba que non hai picos.

Co fin de automatizar este método e, a súa vez facelo máis robusto, modificouse lixeiramente esta parte do algoritmo escollendo outro limiar. O que se fixo foi dividir o dominio en cen anacos iguais, de maneira que, para a meirande parte dos espectros reais, ter unha parte destes anacos sen picos sexa moi probable. En cada anaco calcúlase a media cadrática dos valores do espectro. Posteriormente normalizase a un e tómase a modo de limiar a mediana deste conxunto.

### 3.1.1.1. Filtro de convolución de Savitzky-Golay

Este tipo de filtro dixital é amplamente empregado para o suavizado de sinais contaminadas por ruído. Foi descrito por primeira vez en [2]. Cando se suaviza un sinal estase a asumir que esta varía lentamente con respecto ao intervalo entre puntos adxacentes do dominio e que está contaminado por ruído. Polo xeral é útil substituír cada punto por algún promedio local dos puntos máis próximos. O efecto deste promedio é eliminar as posibles fluctuacións estocásticas dos datos, como por exemplo o posible ruído polo que estean afectados.

Un filtro dixital é aplicado a unha serie de datos equiespazados  $f_i = f(t_i)$  onde  $t_i = t_0 + i\delta$  para algunha constante  $\delta$  e  $i = 0, \dots, N - 1$ . Polo xeral un filtro consistirá na substitución de cada dato no vector de valores discretos  $\vec{f} = (f_0, \dots, f_{N-1})$  por un promedio ponderado dos datos circundantes. Isto é, reempazar  $f_j$  por  $g_j$  onde esta última ven dada por

$$g_j = \sum_{n=-n_L}^{n_R} c_n f_{j+n} \quad \text{con } j = 0, \dots, N-1, \quad (3.3)$$

onde  $n_L$  é o número de puntos que se empregan á dereita do dato  $j$ -ésimo, mentres que  $n_R$  é número de datos empregados a dereita. As constantes  $c_n$  son aquelas que ponderan o promedio.

A idea deste filtro é aproximar a función dentro da ventá (o subdominio que involucra localmente a unha veciñanza de  $n_R + n_L + 1$  valores) por un polinomio, tipicamente cadrático ou cuártico. Para cada punto  $f_j$  realízase un axuste de mínimos cadrados para aproximar os puntos dentro da ventá ao polinomio desexado, entón calcúlase o novo valor  $g_j$  dado polo valor dese polinomio na posición do dato  $f_j$ . Este proceso repítese para tódolos valores do vector de datos  $\vec{f}$ .

A verdadeira vantaxe deste método é que non é preciso realizar un axuste cada vez que se move a ventá. É posible calcular uns coeficientes  $c_n$  tal que a convolución discreta que da ecuación (3.3) proporcione o mesmo resultado.

Para calcular estes coeficientes considérese como se calcula  $g_0$ . Observando a ecuación (3.3) vese que hai que estender os datos a índices negativos, pois o primeiro índice será  $-n_L$ . Suporase que  $f_n$  toma o valor cero para eses valores. En determinado momento ao longo do desenvolvemento aparecerá o vector  $\vec{e}_n$  da base canónica, onde  $n$  tamén tomará valores negativos. Eses serán os vectores unitarios relativos as posicións que van dende  $-n_L$  ata 0. O que hai que facer é axustar por mínimos cadrados un polinomio de grao  $M$  sobre o valor enteiro  $j$  das posicións dos datos, aos valores de  $\vec{f}$  que van dende  $f_{-n_L}$  ata  $f_{n_R}$ . Para isto débense calcular os coeficientes  $\vec{a} = (a_0, \dots, a_M)^T$  do polinomio,  $a_0 + a_1 j + \dots + a_M j^M$ . Unha vez feito isto,  $g_0$  será o valor do polinomio en  $j = 0$ , ou, o que é o mesmo,  $a_0$ . Así pois, nun caso xeral resulta

$$\mathbf{A}\vec{a} = \vec{f}, \quad A_{jl} = j^l, \quad j = -n_L, \dots, n_R, \quad l = 0, \dots, M, \quad (3.4)$$

onde  $\mathbf{A}$  é o que se coñece como matriz de deseño. As ecuacións normais para o vector  $\vec{a}$  en termos de  $\vec{f}$  son

$$(\mathbf{A}^T \mathbf{A})\vec{a} = \mathbf{A}^T \vec{f}, \quad \vec{a} = (\mathbf{A}^T \mathbf{A})^{-1}(\mathbf{A}^T \vec{f}), \quad (3.5)$$

onde

$$\{\mathbf{A}^T \mathbf{A}\}_{jl} = \sum_{k=-n_L}^{n_R} A_{kj} A_{kl} = \sum_{k=-n_L}^{n_R} k^{j+l} \quad (3.6)$$

e

$$\{\mathbf{A}^T \vec{f}\}_l = \sum_{k=-n_L}^{n_R} A_{kl} f_k = \sum_{k=-n_L}^{n_R} k^l f_k. \quad (3.7)$$

Vese entón que

$$\begin{aligned} g_0 = a_0 &= \left\{ (\mathbf{A}^T \mathbf{A})^{-1} \left[ \mathbf{A}^T \sum_{n=-n_L}^{n_R} (\vec{e}_n f_n) \right] \right\}_0 \\ &= \sum_{n=-n_L}^{n_R} \{ (\mathbf{A}^T \mathbf{A})^{-1} (\mathbf{A}^T \vec{e}_n) \}_0 f_n, \end{aligned} \quad (3.8)$$

co cal

$$c_n = \{ (\mathbf{A}^T \mathbf{A})^{-1} (\mathbf{A}^T \vec{e}_n) \}_0 = \sum_{m=0}^M \{ (\mathbf{A}^T \mathbf{A})^{-1} \}_{0m} n^m. \quad (3.9)$$

Supóñase agora que o que interesa non é o sinal en si mesmo senón a súa derivada  $k$ -ésima. Realizando o mesmo axuste, ao tratarse dun polinomio, a derivada é sinxela de calcular no punto  $j$  mediante a expresión  $a_k \cdot k! + a_{k+1} \cdot (k+1)!j + \dots + a_M \cdot (M!/(M-k)!)j^{M-k}$ . Aplicando a mesma lóxica que antes,  $g_0 = a_k \cdot k!$ , de maneira que

$$g_0 = k! \sum_{n=-n_L}^{n_R} \{ (\mathbf{A}^T \mathbf{A})^{-1} (\mathbf{A}^T \vec{e}_n) \}_k f_n. \quad (3.10)$$

En consecuencia, os coeficientes da convolución son

$$c_n = \{ (\mathbf{A}^T \mathbf{A})^{-1} (\mathbf{A}^T \vec{e}_n) \}_k = \sum_{m=0}^M \{ (\mathbf{A}^T \mathbf{A})^{-1} \}_{km} n^m. \quad (3.11)$$

Nótese que o que nos interesa non é o cálculo de  $g_0$ , senón os coeficientes en si mesmos, que nos servirán para calcular calquera  $g_i$  grazas a expresión (3.3). O cálculo de  $g_0$  so é un medio para obter estes coeficientes. O presente método representa un algoritmo moi eficiente e axeitado para calcular numericamente derivadas de orde arbitrario de sinais contaminados con ruído.

### 3.1.2. Reconstrución do espectro

Nesta sección estudarase como caracterizar os picos do espectro  $S(\omega)$  unha vez filtrados. Suporase que é posible aproximalos por unha función de Lorentz, ou Lorentziana  $Y(\omega)$ . En [1] emprégase a seguinte expresión:

$$Y(\omega) = A \frac{\lambda}{\lambda^2 + (\omega - \omega_0)^2}, \quad (3.12)$$

onde  $A$  é a amplitude ou altura,  $\lambda$  é a HWHH (half width at half height ou metade de anchura á metade de altura), magnitude da que xa se falou anteriormente, tamén coñecida como HWHM (half width at half maximum ou metade de anchura á metade do máximo) e  $\omega_0$  é o centro do pico.

O método iterativo seguido en [1] denomínase aproximación proporcional. Primeiramente faise unha estimación dos parámetros do pico  $\lambda$ ,  $A$ , e  $\omega_0$  (o centro é recalculado con maior precisión aínda que o pico xa está localizado). Para iso resólvese o sistema de tres ecuacións e tres incógnitas seguinte.

$$S(\omega_l) = Y(\omega_l) = A \frac{\lambda}{\lambda^2 + (\omega_l - \omega_0)^2}, \quad (3.13)$$

$$S(\omega_m) = Y(\omega_m) = A \frac{\lambda}{\lambda^2 + (\omega_m - \omega_0)^2}, \quad (3.14)$$

$$S(\omega_r) = Y(\omega_r) = A \frac{\lambda}{\lambda^2 + (\omega_r - \omega_0)^2}. \quad (3.15)$$

Como pode verse no sistema anterior, estase a supoñer que o valor de cada Lorentziana coincide co que toma o espectro para os tres puntos característicos da correspondente tripla. Pode consultarse [4] para ver máis detalles acerca da solución deste sistema.

Unha vez obtida a primeira estimación dos parámetros hai que iterar. Sexa  $j \in \{1, 2 \dots J\}$ , sendo  $J$  o número de picos, o índice que recorre os picos, ou triplas.  $x_j$  será a variable que designe os elementos da tripla  $j$ -ésima,  $(\omega_l, \omega_m, \omega_r)_j$ , indistintamente e  $i$  o índice que reflicte en que iteración se está, de tal forma que na iteración  $i$ -ésima o valor da aproximación  $Y_j^i$  do pico  $j$ -ésimo calcúlase como

$$Y_j^{(i)}(x_j) = Y_j^{(i-1)}(x_j) \times \frac{S(x_j)}{\sum_l Y_l^{(i-1)}(x_j)}. \quad (3.16)$$

A iteración deterase cando se incumpra a condición  $Y_l^{(i)}(x_j) \leq S(x_j)$  para algún  $l \in \{1, 2 \dots, J\}$ , é dicir, cando algúnha Lorentziana tome un valor maior

ao do espectro nalgún punto das triplas. En cada iteración hai que recalcular os parámetros resolvendo novamente o sistema de ecuacións (3.13)–(3.15), coa diferenza de que agora non se supón que o valor das curvas Lorentzianas coincide co do espectro, senón que se empregan os valores obtidos por (3.16).

$$Y^{(i)}(\omega_l) = A^{(i)} \frac{\lambda^{(i)}}{(\lambda^{(i)})^2 + (\omega_l - \omega_0^{(i)})^2}, \quad (3.17)$$

$$Y^{(i)}(\omega_m) = A^{(i)} \frac{\lambda^{(i)}}{(\lambda^{(i)})^2 + (\omega_m - \omega_0^{(i)})^2}, \quad (3.18)$$

$$Y^{(i)}(\omega_r) = A^{(i)} \frac{\lambda^{(i)}}{(\lambda^{(i)})^2 + (\omega_r - \omega_0^{(i)})^2}. \quad (3.19)$$

Grazas a (3.17)–(3.19) é posible obter os parámetros da nova curva Lorentziana que se obtén en cada iteración, podendo calcular a seguinte mediante a ecuación (3.16). Nótese que se (3.17)–(3.19) non tivese unha solución analítica e houbera que recorrer a métodos numéricos para resolver o sistema non lineal, todo este proceso resultaría inútil, pois os tempos de cómputo serían moi elevados xa que habería que resolver (3.17)–(3.19) unha vez por iteración. Unha solución adecuada do sistema (3.17)–(3.19) pode atoparse en [4].

## 3.2. Caso bidimensional

Á hora de aproximar a un espectro bidimensional o procedemento a seguir é semellante ao caso dunha soa dimensión, de maneira que é posible diferenciar as mesmas dúas partes do método que se está a aplicar. Por un lado, e nun primeiro momento, débense atopar os picos de radiación e posicionalos correctamente no espectro, por outro, hai que axustalos coas correspondentes funcións analíticas para reconstruír o espectro. Cabe dicir que a xeneralización do procedemento descrito en [1] para espectros bidimensionais é orixinal. Aínda así, en honor a xustiza, hai que dicir que a xeneralización do método é bastante intuitiva.

Para localizar os picos seguirase o mesmo método que foi descrito no apartado anterior para unha dimensión, pero agora, ao ter dúas dimensións, débense empregar derivadas parciais segundas. Cando de filtrar os picos lexítimos de aqueles debidos ao ruído se trate, tamén se empregará unha forma xeneralizada da función puntuación que ten como argumento quintuplas de puntos no canto de triplas.



En canto ao axuste dos picos mediante o método iterativo descrito, neste caso empregaranse funcións de Lorentz bidimensionais:

$$Y(\omega_x, \omega_y) = \frac{A}{1 + ((\omega_x - \omega_{0x})/\lambda_x)^2 + ((\omega_y - \omega_{0y})/\lambda_y)^2}, \quad (3.20)$$

onde  $\omega_x$  e  $\omega_y$  son as coordenadas do espectro,  $\omega_{0x}$  e  $\omega_{0y}$  son as coordenadas do centro da Lorentziana, e  $\lambda_x$  e  $\lambda_y$  son as semianchuras nas correspondentes dimensións.

Unha pequena aclaración sobre a notación se fai precisa neste momento. De agora en diante, cando apareza a variable  $\omega$  con dous subíndices o primeiro deles será o que recorre o eixo horizontal, mentres que o segundo recorrerá eixo vertical, é dicir, o punto  $\omega_{m,n}$  será aquel dado polas coordenadas  $(\omega_m, \omega_n)$ .

### 3.2.1. Identificación dos picos

A continuación verase como se localizan os picos de radiación neste caso de dúas dimensións. Novamente, o primeiro que se vai necesitar é calcular as derivadas segundas pero, como xa se dixo, ao tratar cun dominio bidimensional, este apartado complícase substancialmente. Vaise intentar xeneralizar na medida do posible a lóxica empregada no caso anterior. Para iso o que se fará será localizar a primeira coordenada do pico empregando a derivada parcial segunda con respecto a  $\omega_x$  do espectro e o mesmo para a coordenada  $\omega_y$  coa derivada correspondente. Debido a que as expresións das curvas Lorentzianas teñen unha expresión analítica sinxela, en vez de estudar a matriz hessiana para localizar máximos e mínimos locais, xerárase desta maneira unha lista con tódolos puntos que sexan mínimos locais das dúas derivadas parciais á vez nas súas correspondentes dimensións, é dicir,  $\omega_{m,n}$  será considerado un potencial pico se satisfai a condición

$$\begin{aligned} \frac{\partial^2 S(\omega_{m,n})}{\partial \omega_x^2} &\leq \frac{\partial^2 S(\omega_{m+1,n})}{\partial \omega_x^2} \text{ e } \frac{\partial^2 S(\omega_{m,n})}{\partial \omega_x^2} \leq \frac{\partial^2 S(\omega_{m-1,n})}{\partial \omega_x^2} \text{ e} \\ \frac{\partial^2 S(\omega_{m,n})}{\partial \omega_y^2} &\leq \frac{\partial^2 S(\omega_{m,n+1})}{\partial \omega_y^2} \text{ e } \frac{\partial^2 S(\omega_{m,n})}{\partial \omega_y^2} \leq \frac{\partial^2 S(\omega_{m,n-1})}{\partial \omega_y^2}. \end{aligned} \quad (3.21)$$

Seguindo coa xeneralización do método, agora débese redefinir a función puntuación para que funcione en dúas dimensións, pero para iso hai que trocar as triplas de puntos coas que se traballaba antes, e que eran o argumento

desta función, por quintuplas. Para crear unha quintupla asociaráselle a cada posible pico  $\omega_{m,n}$ , onde agora temos dous subíndices (un por cada dimensión), un intervalo para a variable  $\omega_x$  e outro para a  $\omega_y$  da mesma forma que se facía antes nunha soa dimensión. Novamente definiranse cada un dos intervalos collendo por cada extremo os puntos onde o sinal se anula  $S(\omega_x, \omega_y) = 0$  ou o máximo local máis próximo (destes dous, aquel que sexa máis achegado ao mínimo) da derivada parcial segunda que corresponda, nas coordenadas  $\omega_x$  e  $\omega_y$ , respectivamente.

Destá maneira, por cada mínimo haberá unha quintupla de puntos que veñen dados por:

$$\omega_{m,n} \text{ é tal que } \frac{\partial^2 S(\omega_{m,n})}{\partial \omega_x^2} < 0 \text{ e } \frac{\partial^2 S(\omega_{m-1,n})}{\partial \omega_x^2} > \frac{\partial^2 S(\omega_{m,n})}{\partial \omega_x^2} < \frac{\partial^2 S(\omega_{m+1,n})}{\partial \omega_x^2}$$

$$\text{e } \frac{\partial^2 S(\omega_{m,n})}{\partial \omega_y^2} < 0 \text{ e } \frac{\partial^2 S(\omega_{m,n-1})}{\partial \omega_y^2} > \frac{\partial^2 S(\omega_{m,n})}{\partial \omega_y^2} < \frac{\partial^2 S(\omega_{m,n+1})}{\partial \omega_y^2}$$
(3.22)

$$\omega_{l,n} = \operatorname{argmin}_{0 \leq j < m} \left( \|\omega_{j,n} - \omega_{m,n}\| \left| \left( \overbrace{\frac{\partial^2 S(\omega_{j-1,n})}{\partial \omega_x^2} \geq 0 \text{ e } \frac{\partial^2 S(\omega_{j,n})}{\partial \omega_x^2} < 0}^{\text{Raíz máis próxima}} \right) \right. \right.$$

$$\left. \left. \text{ou } \underbrace{\frac{\partial^2 S(\omega_{j-1,n})}{\partial \omega_x^2} \leq \frac{\partial^2 S(\omega_{j,n})}{\partial \omega_x^2} > \frac{\partial^2 S(\omega_{j+1,n})}{\partial \omega_x^2}}_{\text{Máximo local máis preto}} \right) \right)$$
(3.23)

$$\omega_{r,n} = \operatorname{argmin}_{m < j \leq M} \left( \|\omega_{m,n} - \omega_{j,n}\| \left| \left( \overbrace{\frac{\partial^2 S(\omega_{j+1,n})}{\partial \omega_x^2} \geq 0 \text{ e } \frac{\partial^2 S(\omega_{j,n})}{\partial \omega_x^2} < 0}^{\text{Raíz máis próxima}} \right) \right. \right.$$

$$\left. \left. \text{ou } \underbrace{\frac{\partial^2 S(\omega_{j-1,n})}{\partial \omega_x^2} < \frac{\partial^2 S(\omega_{j,n})}{\partial \omega_x^2} \geq \frac{\partial^2 S(\omega_{j+1,n})}{\partial \omega_x^2}}_{\text{Máximo local máis preto}} \right) \right)$$
(3.24)

$$\omega_{m,d} = \operatorname{argmin}_{0 \leq j < n} \left( \left\| \omega_{m,j} - \omega_{m,n} \right\| \left| \left( \overbrace{\frac{\partial^2 S(\omega_{m,j-1})}{\partial \omega_y^2} \geq 0 \text{ e } \frac{\partial^2 S(\omega_{m,j})}{\partial \omega_y^2} < 0}^{\text{Raiz máis próxima}} \right) \right. \right. \\ \left. \left. \text{ou } \underbrace{\frac{\partial^2 S(\omega_{m,j-1})}{\partial \omega_y^2} \leq \frac{\partial^2 S(\omega_{m,j})}{\partial \omega_y^2} > \frac{\partial^2 S(\omega_{m,j+1})}{\partial \omega_y^2}}_{\text{Máximo local máis preto}} \right) \right) \quad (3.25)$$

$$\omega_{m,u} = \operatorname{argmin}_{n < j \leq N} \left( \left\| \omega_{m,n} - \omega_{m,j} \right\| \left| \left( \overbrace{\frac{\partial^2 S(\omega_{m,j+1})}{\partial \omega_y^2} \geq 0 \text{ e } \frac{\partial^2 S(\omega_{m,j})}{\partial \omega_y^2} < 0}^{\text{Raiz máis próxima}} \right) \right. \right. \\ \left. \left. \text{ou } \underbrace{\frac{\partial^2 S(\omega_{m,j-1})}{\partial \omega_y^2} < \frac{\partial^2 S(\omega_{m,j})}{\partial \omega_y^2} \geq \frac{\partial^2 S(\omega_{m,j+1})}{\partial \omega_y^2}}_{\text{Máximo local máis preto}} \right) \right). \quad (3.26)$$

Así, poderase definir unha quintupla como:

$$\mathbf{p} = (\omega_{l,n}, \omega_{m,n}, \omega_{r,n}, \omega_{m,u}, \omega_{m,d}). \quad (3.27)$$

A función puntuación ven definida por

puntuación( $\mathbf{p}$ ) =

$$\frac{\min \left( \sum_{k=l}^m \left| \frac{\partial^2 S(\omega_{k,n})}{\partial \omega_x^2} \right|, \sum_{k=m}^r \left| \frac{\partial^2 S(\omega_{k,n})}{\partial \omega_x^2} \right|, \sum_{k=d}^n \left| \frac{\partial^2 S(\omega_{m,k})}{\partial \omega_y^2} \right|, \sum_{k=n}^u \left| \frac{\partial^2 S(\omega_{m,k})}{\partial \omega_y^2} \right| \right)}{\max_{\mathbf{q} \in \Omega} \text{puntuación}(\mathbf{q})}. \quad (3.28)$$

Unha vez puntuadas as quintuplas, débense filtrar escollendo un limiar para a puntuación mínima a partir da cal se considerará que son picos. Novamente, o dominio bidimensional será dividido en cen anacos e calcularase a media cadrática dos valores do espectro en cada un deles. Posteriormente, normalizarase a un este conxunto de valores e tomarase como limiar a mediana.

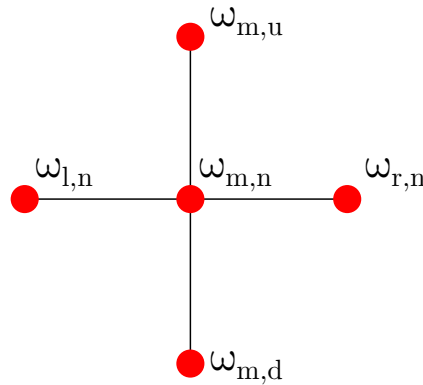


Figura 3.2: Representación esquemática da distribución dos puntos que forman parte dunha quintupla. Os índices  $l$ ,  $r$ ,  $u$  e  $d$  denótanse así polas palabras en inglés *left*, *right*, *up* e *down* (esquerda, dereita, arriba e abaixo).

Para calcular as derivadas en dúas dimensións farase uso do mesmo método que no caso anterior, é dicir, o filtro de convolución de Savitzky-Golay.

### 3.2.1.1. Filtro de convolución de Savitzky-Golay en dúas dimensións

O procedemento calcular os coeficientes para este tipo de filtro dixital é análogo ao caso unidimensional. Neste caso emprégase un polinomio con dúas variables no canto de unha. Para o presente traballo tomáronse os coeficientes calculados en [12] e creouse un programa para realizar a convolución bidimensional que se pode consultar en:

<https://gitlab.com/pablosanchez.aveiga/parametrizacion-de-espectrosen-resonancia-magnetica-nuclear>

Para máis información pódese consultar [5, 6, 7, 12].

### 3.2.2. Reconstrución do espectro

O procedemento para reconstruír o espectro con funcións de Lorentz en dúas dimensións, como as introducidas en (3.20), é moi parecido ao caso unidimensional. O primeiro que se debe facer é calcular unha primeira estimación dos parámetros para cada Lorentziana, é dicir, en cada pico. Unha xeneralización directa do método unidimensional levaría a resolver o seguinte sistema de ecuacións:

$$S(\omega_{m,n}) = Y(\omega_{m,n}) = \frac{A}{1 + ((\omega_m - \omega_{0x})/\lambda_x)^2 + ((\omega_n - \omega_{0y})/\lambda_y)^2}, \quad (3.29)$$

$$S(\omega_{r,n}) = Y(\omega_{r,n}) = \frac{A}{1 + ((\omega_r - \omega_{0x})/\lambda_x)^2 + ((\omega_n - \omega_{0y})/\lambda_y)^2}, \quad (3.30)$$

$$S(\omega_{l,n}) = Y(\omega_{l,n}) = \frac{A}{1 + ((\omega_l - \omega_{0x})/\lambda_x)^2 + ((\omega_n - \omega_{0y})/\lambda_y)^2}, \quad (3.31)$$

$$S(\omega_{m,u}) = Y(\omega_{m,u}) = \frac{A}{1 + ((\omega_m - \omega_{0x})/\lambda_x)^2 + ((\omega_u - \omega_{0y})/\lambda_y)^2}, \quad (3.32)$$

$$S(\omega_{m,d}) = Y(\omega_{m,d}) = \frac{A}{1 + ((\omega_m - \omega_{0x})/\lambda_x)^2 + ((\omega_d - \omega_{0y})/\lambda_y)^2}, \quad (3.33)$$

onde as incógnitas son as semianchuras  $\lambda_x$  e  $\lambda_y$ , a amplitude  $A$ , e as coordenadas do centro da Lorentziana,  $\omega_{0x}$  e  $\omega_{0y}$ . Nótese nas ecuacións (3.29)–(3.33) as variables que aparecen son as coordenadas de frecuencia  $\omega_x$  e  $\omega_y$  dos cinco puntos que conforman a quintupla, isto é:  $\omega_{l,n} = (\omega_l, \omega_n)$ ,  $\omega_{m,n} = (\omega_m, \omega_n)$ ,  $\omega_{r,m} = (\omega_r, \omega_n)$ ,  $\omega_{m,u} = (\omega_m, \omega_u)$  e  $\omega_{m,d} = (\omega_m, \omega_d)$ .

No caso unidimensional existía unha solución para o sistema de tres ecuacións e tres incógnitas que pode atoparse en [4], pero neste caso, aínda que non se pode afirmar que non exista, non foi posible atopala. Ao non ser factible resolver o sistema por métodos numéricos pois, igual que no caso unidimensional, habería que resolvelo unha vez por iteración, o que se fixo foi reduci-lo a dous sistemas de ecuacións análogos a (3.13)–(3.15). Para iso asumírase que a estimación das coordenadas do centro do pico que se obteñen ao identificalos e filtralos, é dicir,  $\omega_{m,n}$ , son correctas. Desta maneira terase  $\omega_m = \omega_{0x}$  e  $\omega_n = \omega_{0y}$ . Así, pódense propoñer os dous sistemas que seguen:

$$S(\omega_{m,n}) = Y(\omega_{m,n}) = \frac{A}{1 + ((\omega_m - \omega_{0x})/\lambda_x)^2}, \quad (3.34)$$

$$S(\omega_{r,n}) = Y(\omega_{r,n}) = \frac{A}{1 + ((\omega_r - \omega_{0x})/\lambda_x)^2}, \quad (3.35)$$

$$S(\omega_{l,n}) = Y(\omega_{l,n}) = \frac{A}{1 + ((\omega_l - \omega_{0x})/\lambda_x)^2}, \quad (3.36)$$

e o segundo sistema de ecuacións,

$$S(\omega_{m,n}) = Y(\omega_{m,n}) = \frac{A}{1 + ((\omega_n - \omega_{0y})/\lambda_y)^2}, \quad (3.37)$$

$$S(\omega_{m,u}) = Y(\omega_{m,u}) = \frac{A}{1 + ((\omega_u - \omega_{0y})/\lambda_y)^2}, \quad (3.38)$$

$$S(\omega_{m,d}) = Y(\omega_{m,d}) = \frac{A}{1 + ((\omega_d - \omega_{0y})/\lambda_y)^2}. \quad (3.39)$$

Formulados así, poderanse resolver do mesmo xeito que (3.13)–(3.15). Nótese que aínda que se supoña que a primeira estimación dos centros é correcta, as coordenadas destes recálculanse en cada iteración, pois na primeira ecuación do sistema dado por (3.34)–(3.36) non se asume que  $\omega_m = \omega_{0x}$ , mentres que en (3.37)–(3.39) o que non se asume é  $\omega_n = \omega_{0y}$ . Para consultar a solución analítica destes sistemas pode verse en [4]. En canto ao valor de  $A$ , que é calculado en ambos sistemas, é preferible tomar o de (3.34)–(3.36), pois a dimensión asociada a  $\omega_x$ , ou dimensión directa, ten maior resolución neste tipo de espectros, o que permite facer unha interpolación máis precisa e que os resultados calculados sexan máis fiables.

Agora que se coñece a primeira aproximación dos parámetros pódese comezar a iterar. Empregarase a mesma notación que no caso unidimensional, de maneira que  $j \in \{1, 2, \dots, J\}$  é o índice que recorre os  $J$  picos,  $x_j$  é a variable que designa os elementos da quintupla  $j$ -ésima,  $(\omega_{m,l}, \omega_{m,n}, \omega_{m,r}, \omega_{u,n}, \omega_{d,n})_j$  e  $i$  o índice que indica a iteración. Hai que ter en conta que os subíndices das frecuencias  $\omega$  fan referencia as coordenadas do dominio, mentres que  $j$  fai referencia aos picos, polo que non están relacionados. O proceso iterativo ven dado pola seguinte relación:

$$Y_j^{(i)}(x_j) = Y_j^{(i-1)}(x_j) \times \frac{S(x_j)}{\sum_{l=1}^J Y_l^{(i-1)}(x_j)}. \quad (3.40)$$

A condición de parada volve a ser a mesma, é dicir, o algoritmo seguirá iterando mentres se cumpra que  $Y_l^{(i)}(x_j) \leq S(x_j)$  para algún  $l \in \{1, 2, \dots, J\}$ . Da mesma forma que no caso anterior, en cada iteración resolverase novamente o sistema de ecuacións, neste caso os dous que veñen dados por (3.34)–(3.36) e (3.37)–(3.39), coa diferenza de que agora non se suporá que as curvas Lorentzianas toman un valor igual ao do espectro, senón aquel que devolve (3.40) en cada iteración. Desta maneira obtéñense os dous sistemas de ecuacións que seguen:

$$Y^{(i)}(\omega_{m,n}) = \frac{A^{(i)}}{1 + ((\omega_m - \omega_{0x}^{(i)})/\lambda_x^{(i)})^2}, \quad (3.41)$$

$$Y^{(i)}(\omega_{r,n}) = \frac{A^{(i)}}{1 + ((\omega_r - \omega_{0x}^{(i)})/\lambda_x^{(i)})^2}, \quad (3.42)$$

$$Y^{(i)}(\omega_{l,n}) = \frac{A^{(i)}}{1 + ((\omega_l - \omega_{0x}^{(i)})/\lambda_x^{(i)})^2}, \quad (3.43)$$

e o segundo sistema de ecuacións dado por

$$Y^{(i)}(\omega_{m,n}) = \frac{A^{(i)}}{1 + ((\omega_n - \omega_{0y}^{(i)})/\lambda_y^{(i)})^2}, \quad (3.44)$$

$$Y^{(i)}(\omega_{m,u}) = \frac{A^{(i)}}{1 + ((\omega_u - \omega_{0y}^{(i)})/\lambda_y^{(i)})^2}, \quad (3.45)$$

$$Y^{(i)}(\omega_{m,d}) = \frac{A^{(i)}}{1 + ((\omega_d - \omega_{0y}^{(i)})/\lambda_y^{(i)})^2}. \quad (3.46)$$

Cos novos parámetros obtidos ao resolver (3.41)–(3.43) e (3.44)–(3.46) pode procederse a unha nova iteración grazas a (3.40) e así sucesivamente ata que se cumpra o criterio de parada.

Na seguinte caixa de texto esquematízase todo o proceso do algoritmo:

#### Algoritmo

1. Identificación de tódolos posibles picos mediante a derivada segunda do espectro.
2. Creación das triplas, no caso unidimensional, ou quintuplas no bidimensional, e almacenado nunha lista.
3. Puntuación das triplas ou quintuplas mediante a función «puntuación».
4. Filtrado dos posibles picos en base ao puntaxe recibido.
5. Primeira estimación dos parámetros das curvas Lorentzianas baixo a suposición de que toman o mesmo valor que o espectro.
6. Iteración recalculando cada vez os novos parámetros.





# Capítulo 4

## Resultados

Unha vez descritos os algoritmos implementados, describíranse os resultados obtidos. O código en C++ do programa, tanto para o caso unidimensional como para o bidimensional, pode atoparse na seguinte ligazón:

<https://gitlab.com/pablosanchez.aveiga/parametrizacion-de-espectros-en-resonancia-magnetica-nuclear>

### 4.1. Caso unidimensional

Comezarase co caso dunha soa dimensión no cal se traballará tanto con espectros simulados artificialmente coma con outros xerados de maneira experimental.

Primeiramente hai que ver que ocorre cun espectro simulado. Para isto empregouse un programa creado a tal propósito que se pode atopar na ligazón dada. Este consiste en dez Lorentzianas distribuídas aleatoriamente polo dominio. A representación gráfica de todas as cantidades realízase con respecto aos índices enteiros dos valores discretos. A integral destas curvas (parámetro  $A$  nas Lorentzianas) é de unha unidade, mentres que o dominio consta de 65536 puntos.  $\lambda$  toma un valor de 0,05 unidades e engadimos ruído gausiano cunha desviación típica de 0,0001 unidades.

O espectro creado con estes parámetros é o que se pode observar na Figura 4.1, mentres que a reconstrución do mesmo devolta polo programa atópase na Figura 4.2.

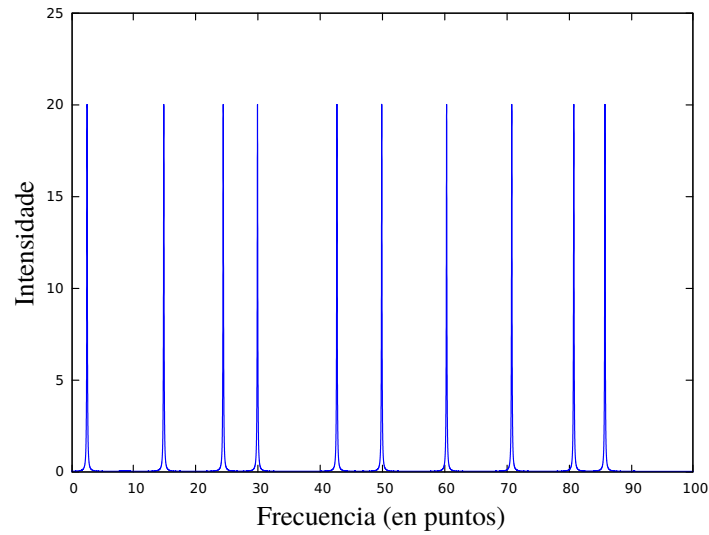


Figura 4.1: Espectro simulado con diez curvas Lorentzianas aleatoriamente distribuidas.

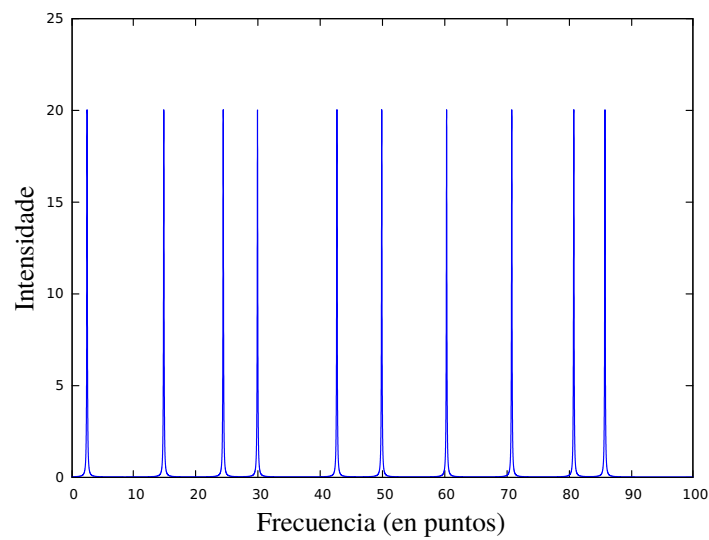


Figura 4.2: Espectro simulado reconstruido mediante el algoritmo a partir de los datos mostrados en la Figura 4.1.

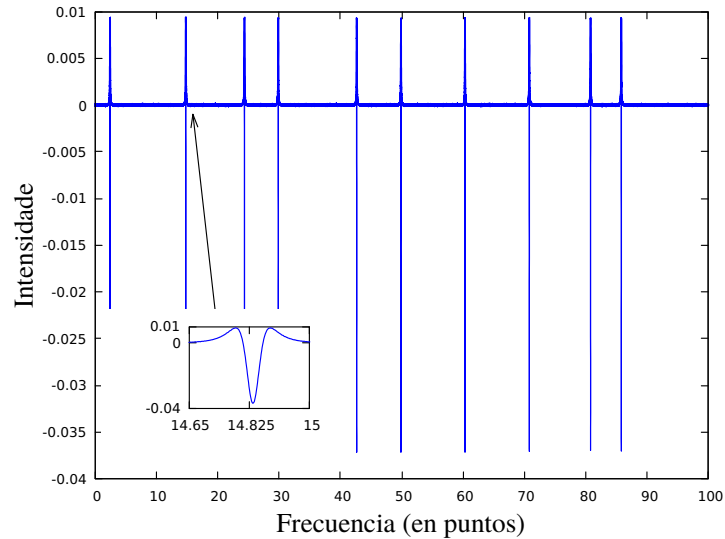


Figura 4.3: Derivada segunda do espectro simulado que se amosa na Figura 4.1 onde pode observarse que a derivada acusa perfectamente a posición dos picos.

Hai que recordar que o cálculo da segunda derivada realizase de xeito diferente ao proposto en [1]. Para ver a implementación do algoritmo de Savitzky-Golay, descrito na sección 3.1.1.1, en C++ pódese consultar o enlace que se da ao principio desta sección. Na Figura 4.3 pódese ver a derivada segunda do espectro simulado. Como se aprecia claramente, a segunda derivada acusa os picos de xeito moi nítido. Desta maneira é posible identificalos con facilidade, pero detectaranse a súa vez tódolos máximos locais do sinal debidos ao ruído. É por iso que agora deben ser puntuados e debidamente filtrados. Para isto empregamos a función *puntuación* definida na ecuación (3.2). Para o espectro simulado obtéñense as puntuacións que se mostran na Figura 4.4. Claramente os dez picos teñen unha puntuación moi superior a aqueles debidos ao ruído.

Agora abordarase un espectro experimental. Na Figura 4.5 móstrase o sinal que é introducido no programa, é dicir, o espectro experimental. Pódese ver que hai tres conxuntos de tres picos que se solapan. Ao seu lado, na Figura 4.6, atópase o espectro reconstruído con funcións Lorentzianas que é devolto polo programa e que, como se ve, reproduce fielmente o experimental.

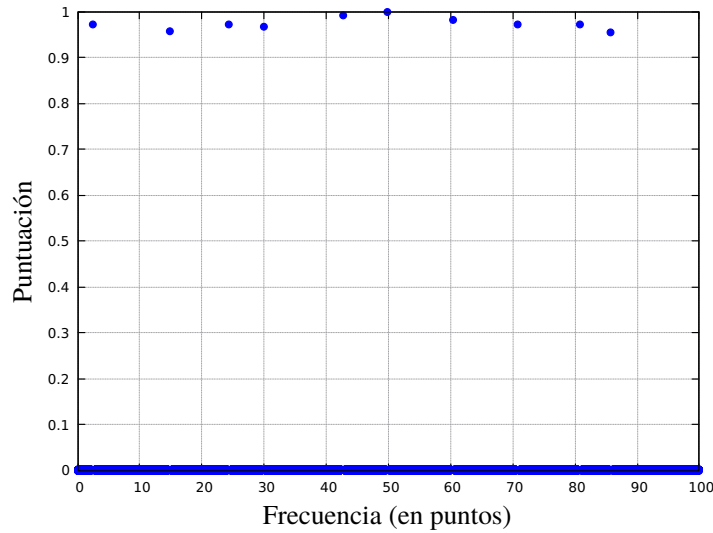


Figura 4.4: Puntuación do espectro simulado da Figura 4.1. A liña azul que se ve abaixo é a acumulación de triplas debido ao ruído con puntuacións próximas a cero.

Hai que ter en conta que neste caso tanto as amplitudes como as anchuras de cada pico son diferentes. Na Figura 4.7 amósase a derivada segunda deste espectro, podendo distinguir os picos que se solapan como mínimos locais da mesma. Isto pode corroborarse observando as puntuacións que se mostran na Figura 4.8. Neste caso vese que os picos teñen puntuacións diferentes en función do pronunciados que sexan pero que, en calquera caso, son moi superiores ás dos picos debido ao ruído. Cabe preguntarse para que rangos de ruído funciona ben o método co que se está a traballar, distinguindo os picos de radiación das perturbacións arbitrarias. Para dar unha resposta a isto o que se fixo foi un estudo paramétrico do número de picos detectados en función do ruído. A maiores variouse tamén o parámetro  $\delta$  para ver como afecta isto á detección de picos. Con este propósito, créanse dous bucles que recorran os parámetros  $\delta$  e  $\sigma$  (desviación típica da distribución normal do ruído gaussiano) respectivamente. En cada iteración simúlase un espectro cos seguintes parámetros  $A = 1$ ,  $\lambda = 0,005$  e dez picos aleatoriamente distribuídos, para posteriormente correr o programa e que identifique os mesmos.

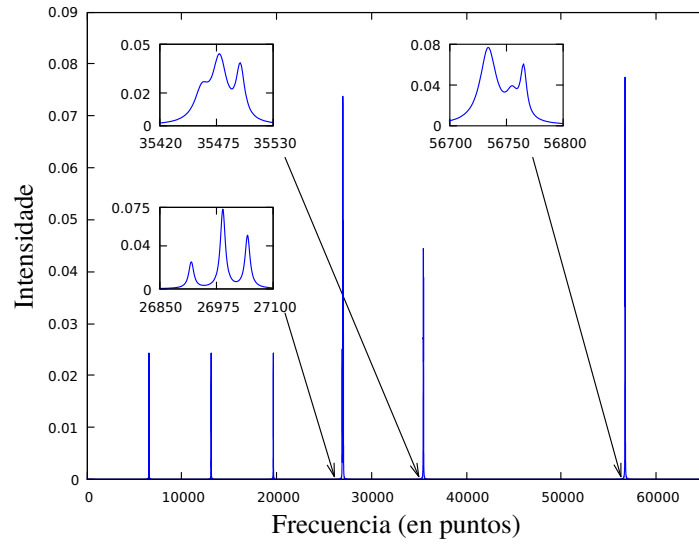


Figura 4.5: Espectro experimental introducido proporcionado pola empresa. Pode observarse que en determinados casos é necesario realizar unha ampliación pois os picos de radiación solápanse entre sí.

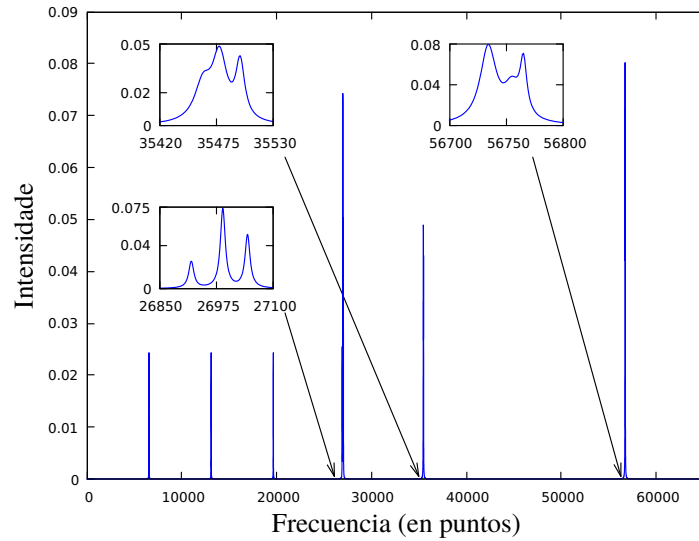


Figura 4.6: Espectro experimental reconstruído a partir dos datos amosados na Figura 4.5 coas necesarias ampliacións.

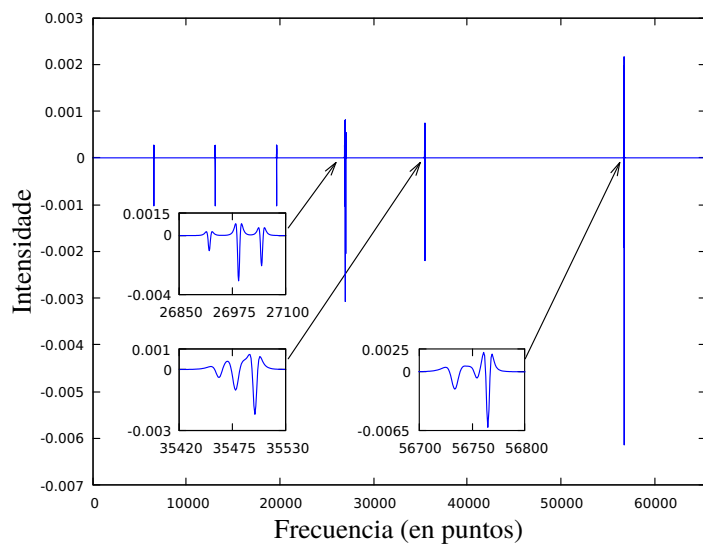


Figura 4.7: Derivada segunda do espectro experimental. As ampliacións desta imaxe mostran as zonas nas que os picos se solapan.

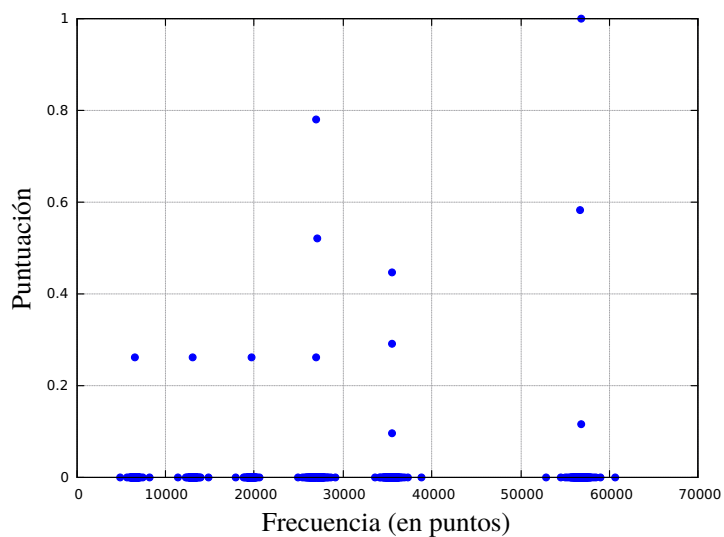


Figura 4.8: Puntuación do espectro experimental onde pode apreciarse que distintas intensidades dos picos conducen a distintas puntuacións.

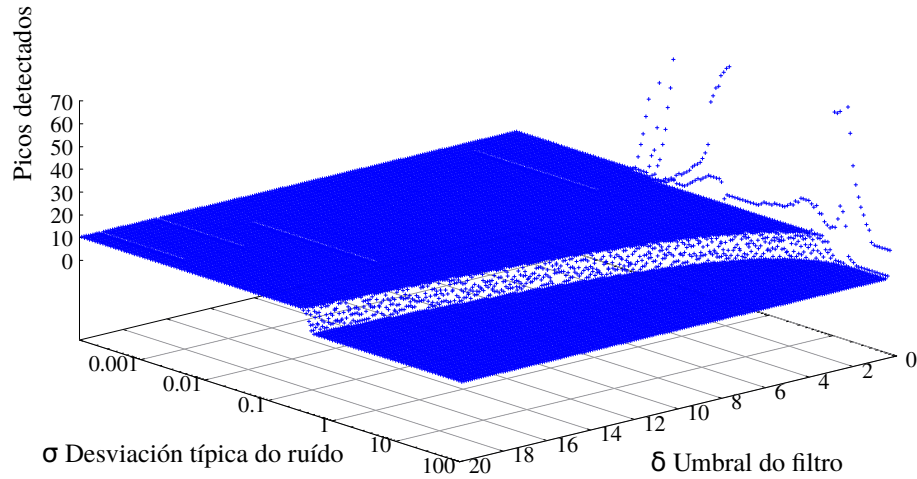


Figura 4.9: Resultado dun experimento que consistiu en simular un espectro con dez picos aleatoriamente distribuídos para distintos valores da desviación típica do ruído  $\sigma$  e do limiar do filtro  $\gamma$  co obxectivo de por a proba as limitacións do algoritmo á hora de detectalos.

O resultado é o que se amosa na Figura 4.9. Como pode verse, na maior parte do dominio detéctanse perfectamente os dez picos, excepto para  $\sigma$  elevados. Compróbase que para  $\delta$  moi baixos, próximos ao cero, o número de picos detectados aumenta drasticamente. Isto débese a que cando se baixa en exceso o valor deste parámetro, apenas se está filtrando o ruído dos picos, detectando as perturbacións debidas a este como picos de radiación. Por outra banda distínguese claramente a transición entre a zona na que se detectan cero picos e na que se detectan os dez.

Existe unha limitación teórica para a resolución deste método. Como se explica en [1], se a distancia entre dous picos é menor que  $\lambda/\sqrt{3}$ , entón a segunda derivada terá un só mínimo local, detectando ambos picos como un. Para estudar este comportamento realizouse un estudo paramétrico en función da separación entre dous picos. A súa vez variouse o parámetro  $\delta$  para ver a súa influencia. Cabe dicir que un dos picos se simulou deliberadamente máis intenso que o outro, sendo, neste caso, o pico 1 o de maior intensidade, cunha amplitude  $A = 1,5$ , mentres que para o pico 2,  $A = 1$ . En canto as súas

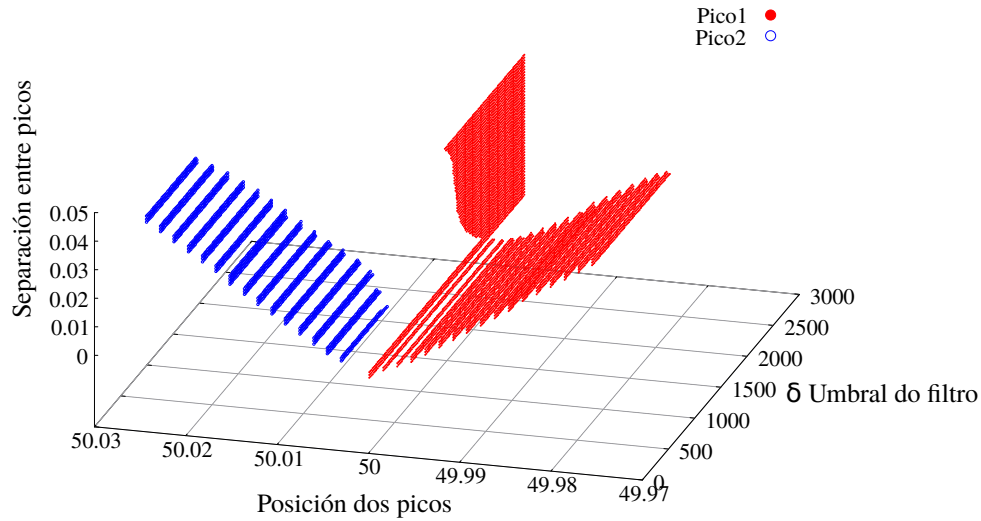


Figura 4.10: Resultado dun experimento que consistía en simular un espectro con dous picos para estudar ata que punto o algoritmo é capaz de detectalos a medida que se solapan. Tamén se variou o limiar do filtro  $\delta$  para estudar como lle afectaba á detección.

anchuras, ambas eran de  $\lambda = 0,005$ . Na Figura 4.10 poden verse os resultados obtidos. Obsérvase como cando se aproximan o suficiente comézase a encontrar só un pico debido a limitación da que se acaba de falar. A separación mínima resoluble teórica é, para este caso, de 0,0029 unidades, mentres que a separación mínima que é posible resolver na práctica é de 0,0053 unidades. Pode afirmarse que ambos valores son bastante próximos.

A discrepancia entre eles débese a que, aínda que para valores inferiores á separación mínima que se atopa na práctica a derivada segunda teña dous mínimos, sendo posible detectar ambos picos, a puntuación que estes teñen é comparable a que teñen as perturbacións provocadas polo ruído, de maneira que ao filtrar os picos só se detecta un.

Ademais, cando se incrementa  $\delta$ , vese que o pico 2, é dicir, o menos intenso, deixa de ser identificado. Este comportamento dáse para un valor de  $\delta \simeq 1000$ . Isto débese a que a súa puntuación é menor que a do pico 1 e ao subir o limiar de puntuacións queda fora do marxe a partir do cal se considera que unha perturbación é pico. Se incrementamos  $\delta$  aínda máis, é localizado un só pico que, independentemente da separación entre o pico 1 e o 2, atópase, aproximadamente, no medio de ambos. O que está a suceder



é que, debido a que o parámetro de filtrado é tan alto, deixase de atopar o pico 1 e o 2 para detectar a suma de ambos coma se fose un só pico.

## 4.2. Caso bidimensional

No que segue amosaranse os resultados obtidos para espectros bidimensionais. Comprobarase que as gráficas son sensiblemente diferentes, pois ao estar en dúas dimensións, e debido a que a cantidade de puntos que constitúen un espectro é moi grande, adoptouse unha representación en forma de mapa de cores, onde a intensidade é representada nunha escala de cor. É necesario facer ampliacións dos lugares nos que se atopan os picos de radiación, pois o seu tamaño é desprezable en comparación coa totalidade do dominio.

Na Figura 4.11 móstrase o espectro experimental que se lle proporciona ao algoritmo. Na Figura 4.12 móstrase a reconstrución do mesmo. Á dereita pode verse a lenda da escala de cor que serve para medir a intensidade de cada punto do espectro. Ademais, dentro do propio espectro vense as distintas ampliacións necesarias para poder distinguir os picos de radiación. Estas están etiquetadas coas letras do A ao F, pero téñase en conta que as etiquetas fan referencia a ampliación, non a picos de radiación, pois hai ampliacións (a D e a E) nas que se poden observar dous picos moi próximos.

Para unha maior claridade, nas Figuras 4.13–4.18, expóñense as ampliacións realizadas sobre o espectro, xunto coas correspondentes reconstrucións, dende unha perspectiva tridimensional. Desta maneira resulta moito máis doado comparar os distintos picos coa función de Lorentz que os modela a simple vista.

Véxase, pois, que sucede coa ampliación *A*, que aparece na Figura 4.13. Como se pode observar, neste caso, o modelo máis adecuado para a perturbación da Figura 4.13a, non semella ser unha función de Lorentz. Isto débese, moi probablemente, a que se está na fronteira do dominio e non aparece o pico completamente debuxado. Pode comprobarse que a reconstrución, na Figura 4.13b, non é precisa.

Outro pico que non é debidamente reconstruído é o que corresponde a ampliación *C*, que se mostra na Figura 4.15. Isto pode deberse a que o modelo que se emprega non se corresponde totalmente coa realidade. Na actualidade

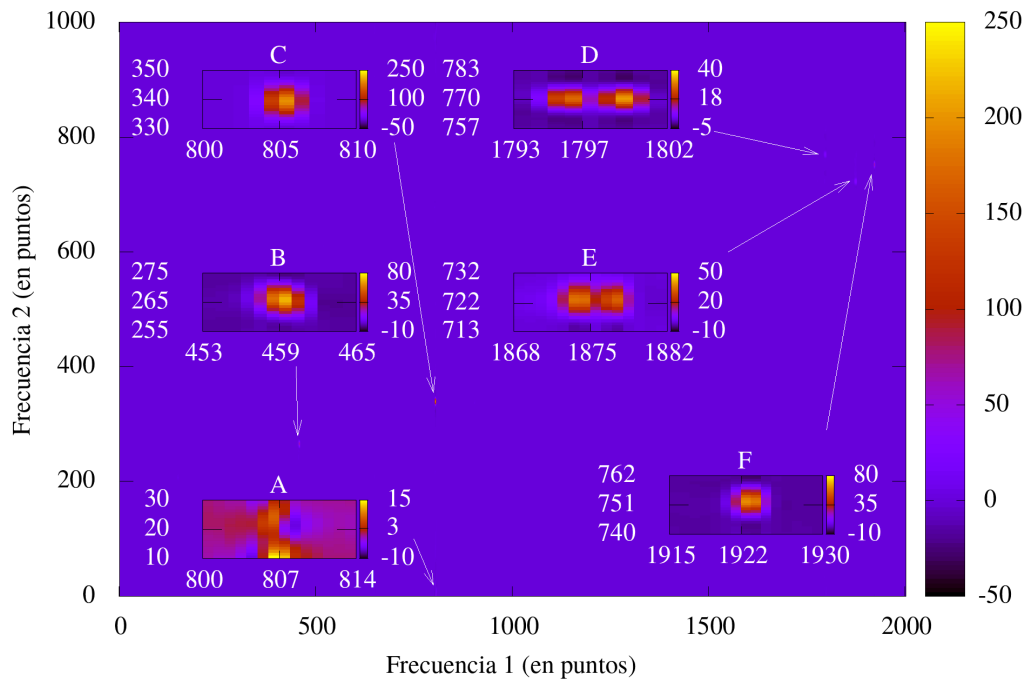


Figura 4.11: Espectro bidimensional proporcionado pola empresa. Para poder visualizar os picos de radiación foi necesario realizar ampliacións, as cales están etiquetadas coas letras A–F.

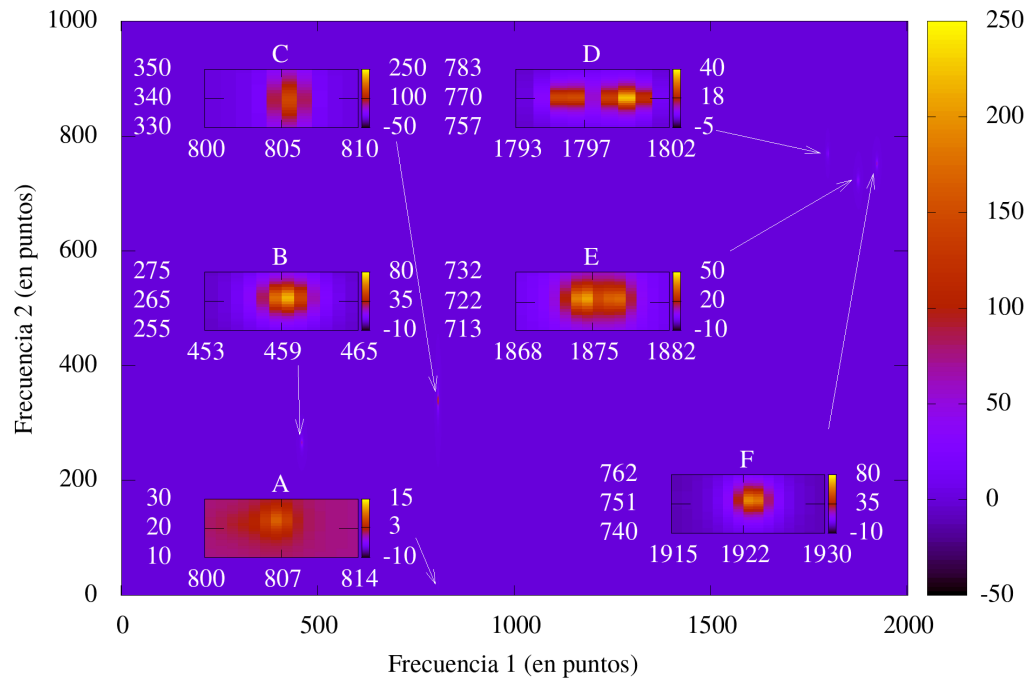


Figura 4.12: Reconstrución do espectro da Figura 4.11. Para poder ver os picos é necesario realizar ampliacións. Estas están etiquetadas coas letras A–F.

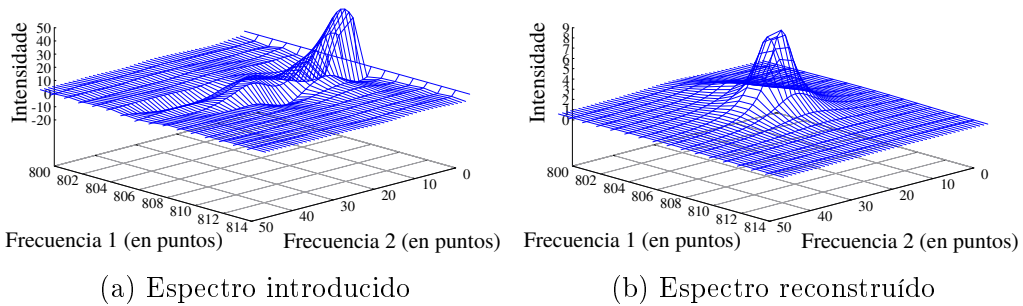


Figura 4.13: Ampliación A: comparación do pico que se atopa na ampliación coa súa reconstrución. Ao tratarse dun pico no bordo do dominio sae cortado e a súa aproximación non é precisa

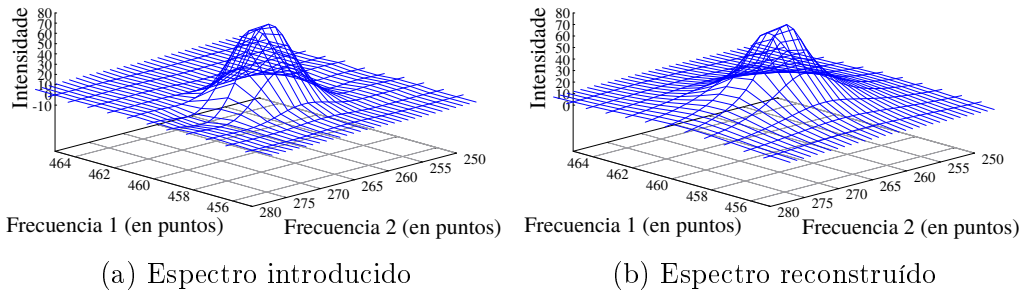


Figura 4.14: Ampliación B: obsérvase a comparación do pico de radiación situado na zona B coa súa reconstrución. A aproximación non presenta problemas.

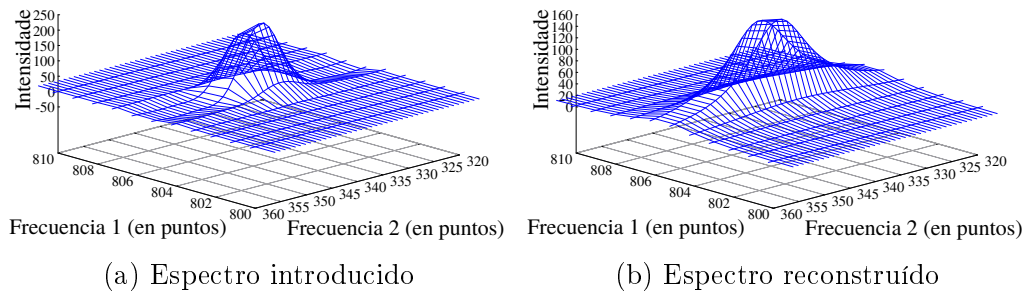


Figura 4.15: Ampliación C: o pico que aparece na zona C é comparado coa súa reconstrución. Obsérvase que a reprodución non é completamente precisa.

utilízanse modelos máis complexos que teñen, a parte da función de Lorentz, compoñentes Gaussianas, entre outras (consúltese [8]). Unha das posibles causas da discrepancia que se atopa na ampliación *C* é a sinxeleza do modelo empregado. Por outra banda, poderían mellorarse os resultados se, unha vez localizados os picos, se reconstrúe un a un individualmente, de maneira que o criterio de parada para a iteración, non dependese de todo o espectro, se non dun pequeno subdominio no que se localiza o pico en cuestión.

Con respecto ao resto de picos a aproximación que proporciona o método é satisfactoria. Incluso nos casos nos que hai dous picos solapándose, como son o das ampliacións *D* e *E*, o método é perfectamente capaz de identificalos

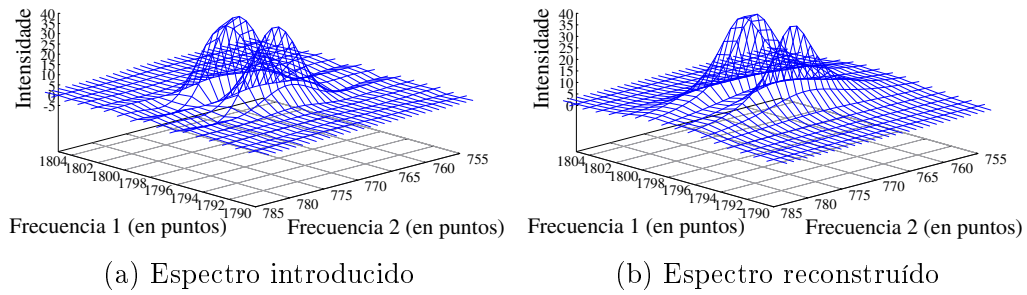


Figura 4.16: Ampliación D: presenta dous picos de radiación que, como se pode observar, son correctamente discriminados e reconstruídos.

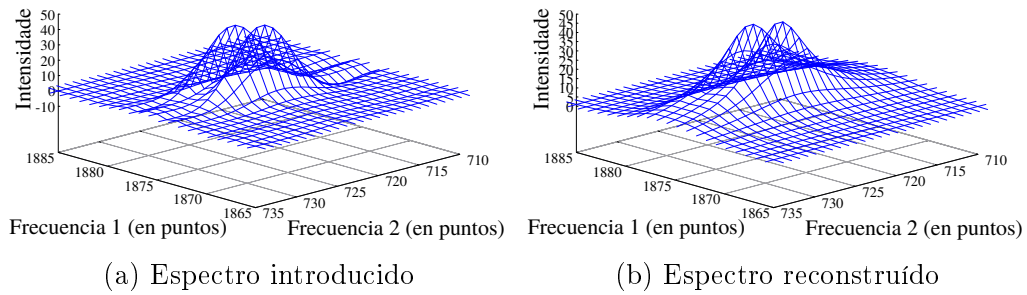


Figura 4.17: Ampliación E: móstranse dous picos de radiación solapados. O método é capaz de distinguilos e reconstruílos sen problema.

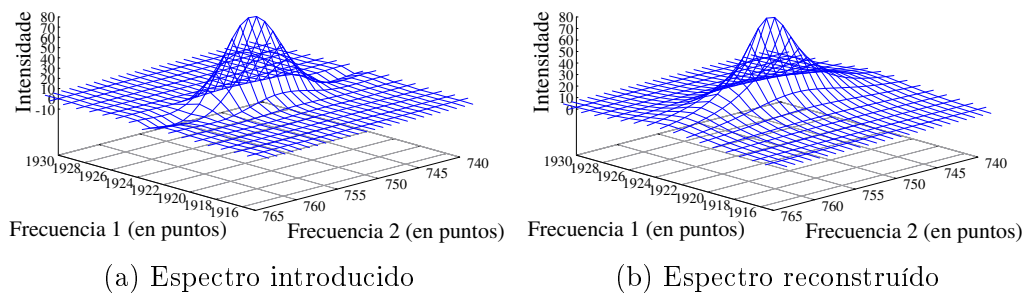


Figura 4.18: Ampliación F: preséntase só un pico de radiación que, como se pode comprobar ao comparalo coa súa reconstrución, se parametriza sen problema.

e reconstruílos. Para comprobar de maneira máis obxectiva a calidade da parametrización do espectro mediante funcións de Lorentz, o que se fixo foi representar nun mapa de cores, análogos aos usados nas Figuras 4.11 e 4.12, a diferenza entre o espectro experimental e o reconstruído en valor absoluto. Na Figura 4.19 pode verse esta representación.

Como se pode observar, na maior parte do dominio, onde non hai picos, a diferenza entre ambos espectros é practicamente nula. Isto é así a pesar de que o espectro experimental está contaminado con ruído incluso nesas zonas. Sucede que as perturbacións debidas ao ruído son desprezables en comparación co erro que se puido cometer en zonas nas que hai picos.

Obviando por un momento os picos que non foron ben parametrizados, é dicir, os correspondentes as ampliacións *A* e *C*, obsérvase un certo patrón no erro que se cometeu co resto de picos. Centrando a atención nas ampliacións *B*, *D*, *E* e *F*, vese que o erro cometido no centro dos picos é moi pequeno, mentres que nun determinado aro arredor do centro é máximo. Isto débese a algo que xa se comentou antes: o modelo empregado non se corresponde totalmente coa realidade, de maneira que as funcións de Lorentz non poden axustar perfectamente os picos de radiación.

En canto a puntuación dos distintos picos, as Figuras 4.20–4.25 mostran a puntuación das zonas onde estes se atopan. Apréciase que, a excepción da zona *A*, onde as puntuacións son máis confusas, os picos distínguense con claridade das perturbacións debidas ao ruído. Incluso nas zonas *D* e *E*, nas que hai dous picos solapados, o método proposto localízalos con precisión.

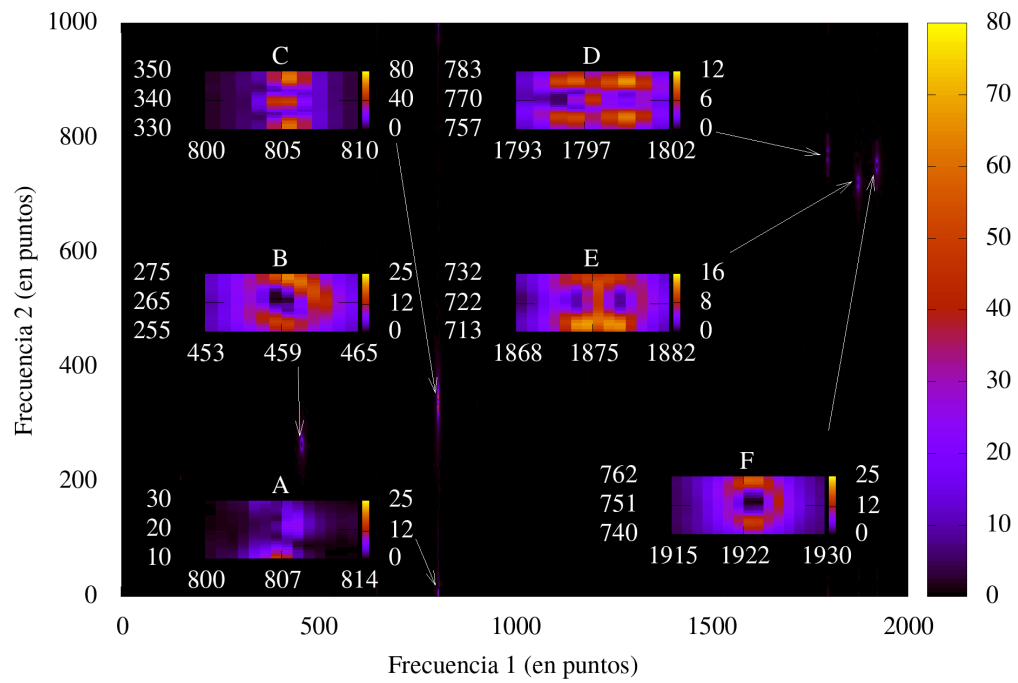


Figura 4.19: Erro absoluto calculado entre o espectro introducido e o reconstruído.

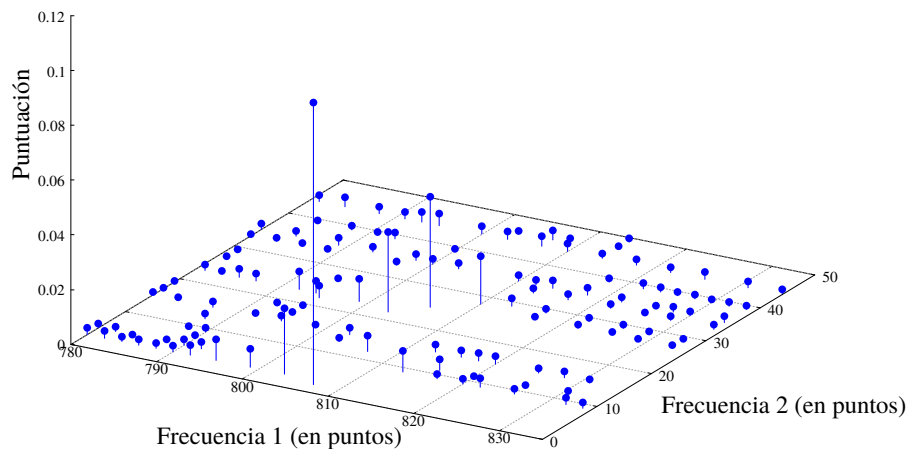


Figura 4.20: Puntuacións calculadas arredor da ampliación A. Neste caso existe certa imprecisión coas puntuacións ao estar preto do bordo.

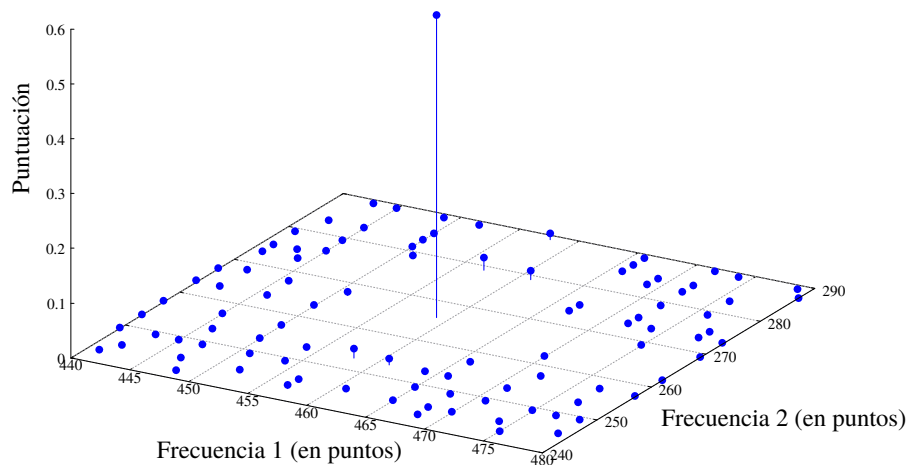


Figura 4.21: Puntuacións calculadas arredor da ampliación B. O pico existente nesta zona diferénciase con claridade.



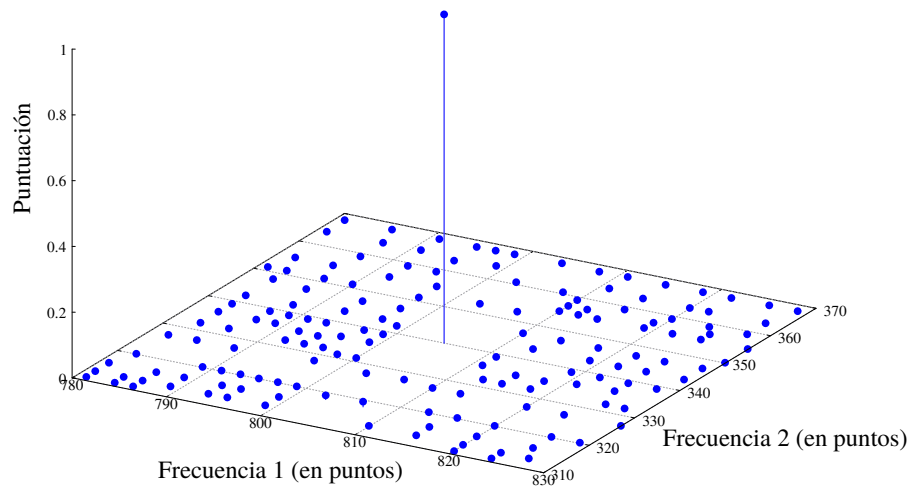


Figura 4.22: Puntuacións calculadas arredor da ampliación C cun único pico de radiación recoñecido.

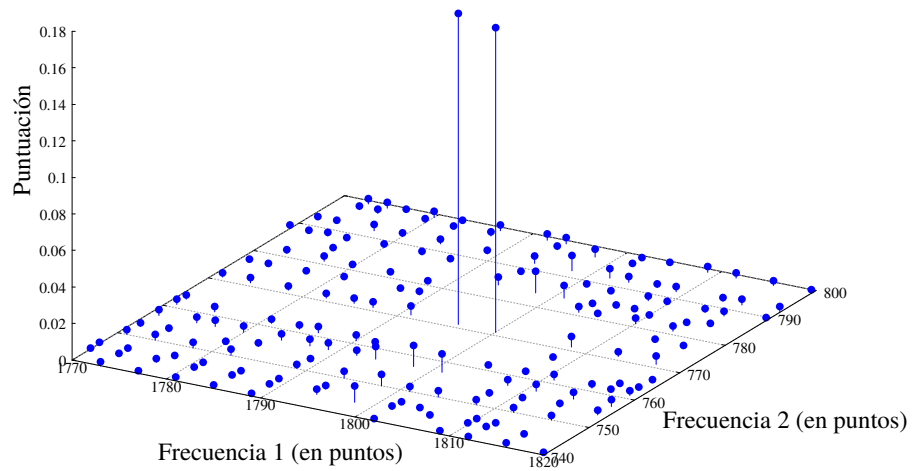


Figura 4.23: Puntuacións calculadas arredor da ampliación D onde existen dous picos que solapan. Obsérvase como o método é capaz de distinguilos de forma precisa.

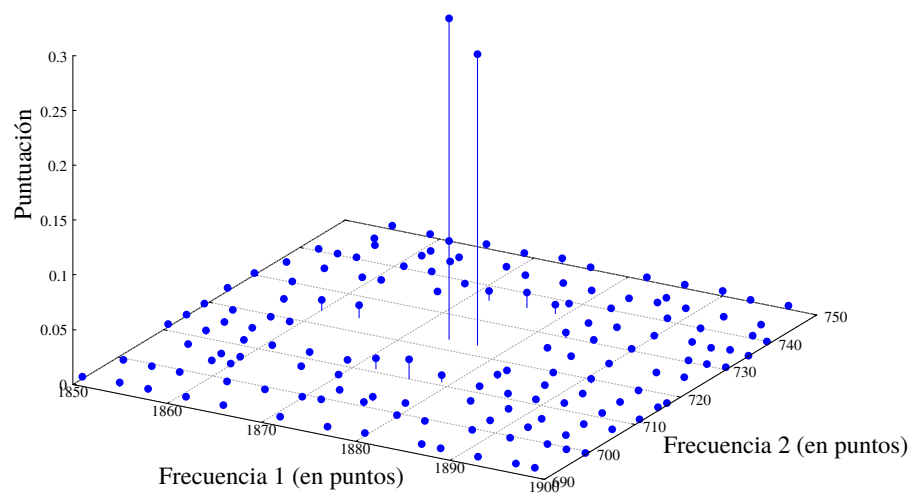


Figura 4.24: Puntuacións calculadas arredor da ampliación E. De novo, dous picos estanse a solapar.

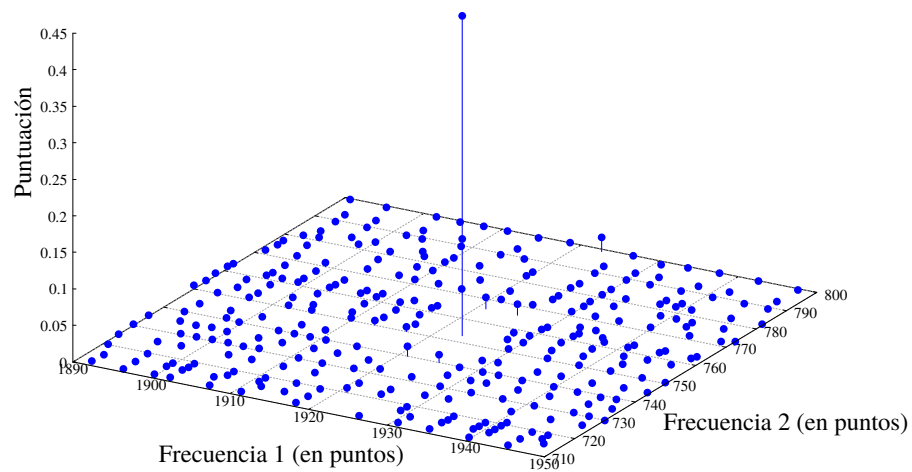


Figura 4.25: Puntuacións calculadas arredor da ampliación F. A puntuación do único pico presente na zona sobresa sobre o resto.

# Capítulo 5

## Conclusións

Ao longo da presente memoria describíronse o método e os resultados numéricos con el obtidos. A modo de conclusión pouco máis se pode facer que valorar os mesmos e decidir se son ou non satisfactorios. Xa se comentou ao longo do traballo que certos aspectos do modelo se poderían revisar e mellorar empregando aproximacións máis complexas para axustar os picos de radiación. Intentarase aquí facer unha pequena aclaración sobre estes aspectos.

O caso unidimensional é máis sinxelo, tanto a hora de desenvolver o algoritmo e programalo, como en canto á análise dos resultados numéricos que se obtiveron ao aplicalo. Hai que sinalar que, aínda que a xeneralización a dúas dimensións pode parecer intuitiva (véxase por exemplo [6], [3] ou [12]), na práctica xorden problemas pois o código complícase notablemente.

Pódese dicir que a primeira parte do algoritmo, é dicir, a localización e puntuación dos picos, é moi satisfactoria. O método de detectar os picos empregando a segunda derivada para posteriormente filtralos a través da función puntuación é computacionalmente superior a outros que se veñen empregando e que, en moitos casos, son algoritmos de recoñecemento de imaxe para buscar bordos.

Esta primeira parte non presentou dificultades tampouco en dúas dimensións, polo menos en canto aos resultados numéricos se refire. Pódese argumentar que no caso da ampliación  $A$  que se mostra na Figura 4.13 o método non funciona como é debido e que as puntuacións expostas na Figura 4.20 son confusas e pouco clarificadoras, pero como xa se sinalou na Sección 4.2, neste caso os problemas debíanse a que se está a tratar cun pico que está no bordo do espectro, de maneira que sae cortado. Aínda así, o principal erro

neste caso ven a hora de reconstruír o pico, non ao detectalo.

Unha pequena aclaración con respecto ao artigo [1] é necesaria. Na parte da reconstrución dos picos, dentro da iteración, hai que resolver un sistema de tres ecuacións con tres incógnitas, como xa se apuntou na correspondente sección. Nun dos apéndices de [1] aparece a solución explícita deste sistema, pero por máis que se intentou implementar esa solución, parecía estar errada. Debido a isto foi consultada a bibliografía que aparecía [1], atopando que a solución deste sistema fora consultada en [4]. Nese artigo, que trata da interpolación de espectros no dominio das frecuencias mediante funcións de Lorentz, pode atoparse unha solución explícita correcta, que foi a que finalmente se implementou.

Como xa se comentou na Sección 3.2.2, unha xeneralización do método unidimensional canónica pasaría pola resolución dun sistema non lineal de cinco ecuacións con cinco incógnitas. Isto non foi posible debido a que a solución explícita dun sistema non lineal deste tamaño non é manexable para a súa implementación. Evidentemente, podería empregarse un método numérico para resolver este sistema non lineal, pero, como xa foi aclarado, ao ter que resolver o sistema unha vez por iteración, este procedemento incrementaría sensiblemente o tempo de cálculo. De todos xeitos, foi posible reducir o sistema non lineal ao do caso unidimensional salvando así esta dificultade computacional.

Por último descríbense posibles formas de mellorar o traballo realizado. A principal mellora que se debería aplicar é cambiar o modelo dos picos de radiación. No caso unidimensional as limitacións do que se ven empregando fanse máis evidentes. Na Figura 4.19 obsérvase claramente que o modelo é insuficiente pois, aínda que aproxima ben os picos no centro, pode apreciarse que o erro aumenta ao afastarse do mesmo.

# Bibliografía

- [1] HYUNG-WON KOH, SASIDHAR MADDULA, JÖRG LAMBERT, ROLAND HERGENRÖDER e LARS HILDEBRAND, «An approach to automated frequency-domain feature extraction in nuclear magnetic resonance spectroscopy», *Journal of Magnetic Resonance*, **201**, páx. 146–156, 2009.
- [2] ABRAHAM SAVITZKY e MARCEL J. E. GOLAY, «Smoothing and Differentiation of Data by Simplified Least Squares Procedures», *Analytical Chemistry*, **36**, páx. 1627–1639, 1964.
- [3] W. P. AUE, E. BARTHOLDI e R. R. ERNST, «Two-dimensional spectroscopy. Application to nuclear magnetic resonance», *The Journal of Chemical Physics*, **Volume 64**, **Número 5**, páx. 2229–2246, 1976.
- [4] CARLO GIANCASPRO e MELVIN B. COMISAROW, «Exact Interpolation of Fourier Transform Spectra», *Applied Spectroscopy*, **Volume 37**, **Número 2**, páx. 153–166, 1983.
- [5] DAVID J THORNLEY, «Anisotropic Multidimensional Savitzky Golay kernels for Smoothing, Differentiation and Reconstruction», (Technical report). Imperial College Department of Computing. 2066/8.
- [6] KENNETH L. RATZLAFF e JEAN T. JOHNSON, «Computation of two-dimensional polynomial least-squares convolution smoothing integers», *Analytical Chemistry*, **Volume 61**, páx. 1303–1305, 1989.
- [7] WILLIAM H. PRESS, SAUL A. TEUKOLSKY, WILLIAM T. VETTERLING e BRIAN P. FLANNERY, *Numerical Recipes, The Art of Scientific Computing*, terceira edición, Cambridge University Press, 2007.
- [8] JEFFREY C. HOCH e ALAN S. STERN, *NMR Data Processing*, primeira edición, Wiley-Liss, Inc., 1996.

- [9] NICOLAI M. JOSUTTIS, *The C++ Standard Library, A Tutorial and Reference*, primeira edición, Addison Wesley, 1999.
- [10] MALCOLM H. LEVITT, *Spin Dynamics, Basics of Nuclear Magnetic Resonance*, segunda edición, John Wiley & Sons, Ltd., 2008.
- [11] JAMES KEELER, *Understanding NMR Spectroscopy*, segunda edición, John Wiley & Sons, Ltd., 2010.
- [12] JOHN KRUMM, *Savitzky-Golay Filters for 2D Images*, Microsoft Research, Microsoft Corporation, 2001.
- [13] ALJOHANI, HIND. *Role of Carboxylate ligands in the Synthesis of AuNPs: Size Control, Molecular Interaction and Catalytic Activity.*, 2016