

UNIVERSIDAD NACIONAL DE LA PLATA  
FACULTAD DE CIENCIAS EXACTAS  
DEPARTAMENTO DE FÍSICA

ANÁLISIS DE SEÑALES E  
INFERENCIA ESTADÍSTICA

TESIS DOCTORAL

LAURA REBOLLO NEIRA

1986

**UNIVERSIDAD NACIONAL DE LA PLATA**

**FACULTAD DE CIENCIAS EXACTAS**

**DEPARTAMENTO DE FISICA**

**TESIS DOCTORAL**

**análisis de señales e inferencia estadística**

**LAURA REBOLLO NEIRA**

**ANGEL PLASTINO**

**DIRECTOR**

**1986**

*a los compañeros desaparecidos  
y a sus madres, las nuestras,  
las de Plaza de Mayo.*

# INDICE

INTRODUCCION	1
CAPITULO I: TEORIA DE LA INFORMACION Y MECANICA ESTADISTICA	5
I.1. El Operador Estadístico y su Ignorancia	7
I.2. Construcción del Operador Estadístico Optimo	11
I.3. Resultados Clásicos	18
I.4. Conclusiones	20
Referencias	23
CAPITULO II: REPRESENTACION DE SEÑALES. METODO DE INVERSION	24
II.1. Efecto de Muestreo	26
II.2. Superposición de Señales	32
II.3. Ignorancia Asociada a la Representación de Señales	36
II.4. Método de Inversión	39
II.5. Convergencia dentro de la Información Disponible	42
II.6. Conclusiones	44
Referencias	47
CAPITULO III: ALGUNAS APLICACIONES	48
III.1. Inversión a partir de Medidas Puntuales	49

<b>III.2.</b>	<b>Inversión a partir de los Momentos de una Distribución</b>	<b>51</b>
<b>III.3.</b>	<b>Ejemplo Numérico</b>	<b>52</b>
<b>III.4.</b>	<b>Deconvolución del Ancho Instrumental</b>	<b>55</b>
<b>III.4.1.</b>	<b>Simulación</b>	<b>55</b>
<b>III.4.2.</b>	<b>Caso Realista: Deconvolución de un Perfil producido por Difracción de Rayos X</b>	<b>57</b>
<b>III.5.</b>	<b>Distribución de Tamaños Cristalinos en Arcillas</b>	<b>62</b>
<b>III.6.</b>	<b>Conclusiones</b>	<b>65</b>
	<b>Referencias</b>	<b>66</b>
<b>CAPITULO IV:</b>	<b>SENALES ESTADISTICAS</b>	<b>67</b>
<b>IV.1.</b>	<b>Definición de Señales Estadísticas</b>	<b>68</b>
<b>IV.2.</b>	<b>El Problema del Interestratificado</b>	<b>71</b>
<b>IV.3.</b>	<b>Tratamiento Semiclásico de la Dispersión de Rayos X</b>	<b>73</b>
<b>IV.4.</b>	<b>Dispersión por una capa</b>	<b>77</b>
<b>IV.5.</b>	<b>Difracción de N Capas Interestratificadas</b>	<b>80</b>
<b>IV.6.</b>	<b>Señales Respuesta de Sistemas Interestratificados en el Formalismo Número de Ocupaciones</b>	<b>83</b>
<b>IV.7.</b>	<b>Medidas como Observables en el Formalismo Número de Ocupaciones</b>	<b>86</b>
<b>IV.8.</b>	<b>Conclusiones</b>	<b>92</b>
	<b>Referencias</b>	<b>93</b>

<b>CAPITULO V: ANALISIS DEL DIAGRAMA DE DIFRACCION DE ARCILLAS INTERESTRATIFICADAS</b>	<b>94</b>
<b>V.1. Características Generales de los dos Componentes de una Muestra de Ilita-Montmorillonita</b>	<b>95</b>
<b>V.2. Determinación del Grado de Interestratificación de dos Muestras Diferentes de Ilita-Montmorillonita</b>	<b>98</b>
<b>V.3. Estimación de Errores</b>	<b>105</b>
<b>V.4. Conclusiones</b>	<b>107</b>
Referencias	109
<b>CONCLUSIONES GENERALES</b>	<b>110</b>
<b>AGRADECIMIENTOS</b>	<b>112</b>

## INTRODUCCION

Comúnmente se denomina señal a todo lo que de algún modo posee la capacidad de transportar información. Aquí nos ocuparemos de señales que contengan información acerca del estado de sistemas físicos, y por lo tanto las asumiremos interviniendo en la observación de los mismos.

Los instrumentos de medida procesan señales transformando los datos en alguna forma reconocida para el observador. Es por esto que existe una amplia variedad de formas de señal que aparecen, incluso dentro de un sistema de proceso particular<sup>1</sup>.

Para independizar a la señal de su procesamiento, adoptaremos una representación vectorial, considerándola como elemento de un espacio de señales<sup>2</sup>. Supondremos que este espacio de señales es un espacio de Hilbert y llamaremos ket ( $| \rangle$ ) al vector que representa una señal.

La información que estos kets contienen, puede extraerse realizando medidas sobre los mismos. Por lo tanto, caracterizaremos las señales estudiadas diciendo que son

vectores que contienen información acerca del estado de sistemas físicos y sobre los cuales es posible efectuar aplicaciones que los conviertan en valores reales. Dichas aplicaciones se denominan funcionales y las utilizaremos para representar medidas sobre señales, limitándonos a considerar sólo el caso de medidas que admitan ser expresadas como funcionales lineales.

Cuando se trata con algún tipo particular de señales, frecuentemente se considera que las mismas forman un conjunto, definido por medio de una propiedad que debe cumplir todo elemento que pertenezca a dicho conjunto. Tal propiedad crea un subconjunto de todas las señales posibles y si es suficientemente restrictiva, el subconjunto que genera puede ser manejado más fácilmente que un conjunto no restrictivo. La elección de la propiedad debe hacerse respecto del problema en cuestión. Para poder definir el conjunto de señales de las que nos ocuparemos, diremos que una señal es representable en alguna base de señales  $\{|r_n\rangle\}$  si admite la siguiente expansión:

$$|f\rangle = \sum_n C_n |r_n\rangle$$

Llamaremos representación de la señal respecto de la base  $\{|r_n\rangle\}$ , al conjunto de coeficientes  $\{C_n\}$ .

La propiedad común a todas las señales que analizaremos, es que cualquiera de ellas admite una representación positiva en alguna base. Por lo tanto definimos el conjunto



$S_{\geq}$  de señales de representación positiva como:

$$S_{\geq} = \{ |f\rangle ; \text{ existe una base donde } |f\rangle = \sum_n C_n |r_n\rangle$$

con  $C_n \geq 0$  para todo  $n$

La definición puede extenderse al caso continuo:

$$S_{\geq} = \{ |f\rangle ; \text{ existe una base donde } |f\rangle = \int g(\alpha) |\alpha\rangle d\alpha$$

con  $g(\alpha) \geq 0$  para todo  $\alpha$

El interés por el análisis de este tipo de señales, se debe a que en su representación está contenida la estadística del sistema.

Puesto que intentar dar la representación de una señal, a partir de un número finito de medidas sobre la misma, es un problema que en general admite más de una solución, adoptaremos el criterio que provee la Teoría de la Información para elegir una de las posibles soluciones. Aceptando el postulado de la teoría, a tal solución la presentamos como la mejor que podemos dar en base a la información conocida. Comprobamos la validez de tal afirmación a través del propio conjunto de medidas disponibles.

En la primera parte de este trabajo, supondremos conocida la base donde una señal admite representación positiva. Para tratar estos casos, presentamos un formalismo y desarrollamos una metodología que permite instrumentar el criterio que adoptamos. Ejemplificaremos la bondad del

método aplicándolo a la resolución de ejemplos numéricos y situaciones realistas.

A partir del capítulo IV, discutiremos la forma de elegir una base adecuada, en término del tipo de información que contiene una representación. De todas las señales pertenecientes al conjunto  $S_n$ , seleccionaremos aquellas que contienen información acerca de la distribución de estados de un sistema y las llamaremos "señales estadísticas". Por medio del método previamente presentado, encontraremos la distribución de estados, partiendo de medidas sobre estas señales estadísticas. Ilustraremos el procedimiento propuesto, estudiando sistemas de arcillas interestratificadas, por medio del diagrama de difracción de rayos X que produce.

#### REFERENCIAS

1. L.E. FRANKS, Teoría de la señal, REVERTE (1975)
2. C.E. SHANNON, Communication in the presence of noise, Proc. IRE, vol. 37, N° 1, (1947)

## CAPITULO I

### TEORIA DE LA INFORMACION Y MECANICA ESTADISTICA

En 1948 Claude E. Shannon<sup>1</sup> propuso y estudió un modelo matemático sobre sistemas de comunicación, llamado TEORIA DE LA INFORMACION (TI). Sus ideas se extendieron y aplicaron rápidamente en otras disciplinas, particularmente dos trabajos desarrollados por E.T. Jaynes<sup>2,3</sup> en 1957 muestran cómo toda la mecánica estadística puede derivarse a partir de la TI en forma sencilla.

La propuesta de Shannon consiste en cuantificar la ignorancia que representa disponer de una distribución de probabilidades  $\{p_i\}$  definidas sobre un conjunto de  $M$  eventos mutuamente excluyentes, mediante la función falta de información dada por:

$$I = -K \sum_{i=1}^N p_i \ln p_i$$

donde  $K$  es la unidad de información. La función  $I$  es la herramienta con la que se trabaja a efectos de definir y

operar matemáticamente con el concepto de información.

En el presente capítulo nos dedicaremos a mostrar la conexión entre la TI y la mecánica estadística. La deducción se hará sobre el caso cuántico y al final del capítulo se resumirán los resultados que surgen de un tratamiento clásico. Para mayor profundización sobre estos temas citamos los trabajos de Jaynes<sup>2,3</sup> y las obras de A. Katz<sup>4</sup> y A. Hobson<sup>5</sup>

## I.1 EL OPERADOR ESTADISTICO Y SU "IGNORANCIA"

El conocimiento total de un sistema mecánico cuántico equivale a haber determinado exactamente el vector de estado que lo describe. Si esto no es posible, como en general sucede, consideraremos el caso en que a diferentes estados se atribuyen diferentes probabilidades. Al hacer esta asignación de probabilidades debemos asegurarnos de que los estados sean mutuamente excluyentes si pretendemos que las probabilidades estén normalizadas a la unidad. Por lo tanto sólo podremos atribuir probabilidades a estados que forman una base ortonormal y completa del espacio de Hilbert.

Consideremos una base de estas características  $\{|i\rangle\}$  y asignemos probabilidades  $p_i$  a cada uno de sus elementos. Podemos entonces construir un operador estadístico del siguiente modo:

$$\hat{P} = \sum_i |i\rangle p_i \langle i| \quad (I.1)$$

$\hat{P}$  es un operador de probabilidad. Cada estado con probabilidad bien definida es autoestado de  $\hat{P}$ , siendo su probabilidad el correspondiente autovalor.

La utilidad del operador estadístico reside en que éste permite calcular el promedio estadístico de cualquier observable  $\hat{O}$ :

$$\langle \hat{O} \rangle = \sum_i p_i \langle i | \hat{O} | i \rangle = \text{TR}(\hat{P} \hat{O}) \quad (I.2)$$

De la definición misma de  $\hat{P}$  se pueden inferir dos

propiedades básicas:

1. La normalización de las probabilidades,  $\sum_i p_i = 1$ , se puede expresar mediante:

$$\text{TR}(\hat{P}) = 1 \quad (I.3)$$

2. Como todos los autovalores  $p_i$  de  $\hat{P}$  son no negativos, éste es un operador definido positivo:

$$\hat{P} \geq 0 \quad (I.4)$$

Además puede demostrarse que  $\hat{P}$  se transforma canónicamente como cualquier otro operador hermitico. Es decir, si  $\hat{W}$  es el generador de una transformación canónica arbitraria y  $t$  su parámetro asociado, entonces:

$$d\hat{P}/dt = i/\hbar [\hat{W}, \hat{P}] \quad (I.5)$$

Al construir el operador estadístico hemos supuesto que poseemos cierto conocimiento parcial que nos permite asociar a cada estado  $|i\rangle$  de una base ortonormal y completa cierta probabilidad  $p_i$ .

El conocimiento total del estado del sistema corresponde al caso en que todas las  $p_i$  son nulas, excepto una, digamos  $p_j$ , que vale uno e indica que el sistema se encuentra en el estado puro  $|j\rangle$ . Cualquier otra distribución de probabilidad implica una cierta ignorancia respecto del estado del sistema que llamaremos I:

$$I = -K \sum_i p_i \ln p_i \quad (1.6)$$

donde K es una constante positiva. Esta ecuación puede llevarse a la forma:

$$I = -K \text{TR}(\hat{P} \ln \hat{P}) \quad (1.7)$$

ya que  $\hat{P}$  es diagonal en la base  $|i\rangle$ :

$$\begin{aligned} \text{TR}(\hat{P} \ln \hat{P}) &= \text{TR}\left(\sum_i |i\rangle p_i \ln p_i \langle i|\right) \\ &= \sum_i p_i \ln p_i \end{aligned}$$

Aquí vemos que el operador estadístico no es otra cosa que la matriz densidad y si hacemos  $K=K_B$  (constante de Boltzman) su ignorancia asociada I es la entropía S introducida por Von Neumann\*:

$$S = -K_B \text{TR}(\hat{P} \ln \hat{P}) \quad (1.8)$$

Desde el punto de vista de la TI, la entropía S es una medida de cuánto desconoce el observador el estado en que el sistema se encuentra. Si la información es completa, es decir, si el sistema considerado está en un estado puro, entonces su entropía es nula. En caso contrario,  $S > 0$  y será tanto mayor cuanto más grande sea el número de estados con probabilidad no nula y más homogénea sea la distribución de probabilidad.

Si no poseemos ninguna información acerca del estado del

sistema, es decir si nuestra ignorancia respecto del estado del sistema es la máxima posible, la distribución de probabilidad que debemos elegir debe ser tal que  $I$  también sea máxima. Esto conduce al siguiente primer principio básico:

"A priori, todos los estados de una base ortonormal y completa son equiprobables" ya que es la distribución uniforme la que hace adoptar a  $I$  su máximo valor.



## I.2 CONSTRUCCION DEL OPERADOR ESTADISTICO OPTIMO

Dado que el operador estadístico  $\hat{P}$  describe el conocimiento que poseemos del sistema, éste será útil en los casos en que sea imposible determinar completamente el estado del sistema, pero se cuente con información parcial acerca del mismo.

En general, esta información estará dada en la forma de valores de expectación de ciertos observables conocidos  $\hat{O}_i$ :

$$\langle \hat{O}_i \rangle = o_i, \quad i=1, \dots, M \quad (I.9)$$

Al construir el operador estadístico debemos incluir en él la información que estas ecuaciones contienen, es decir debemos imponerle a  $\hat{P}$  que reproduzca los valores medidos  $o_i$ :

$$\text{TR}(\hat{P} \hat{O}_i) = o_i, \quad i=1, \dots, M \quad (I.10)$$

Sin embargo las condiciones impuestas sobre  $\hat{P}$  no son aún suficientes como para determinarlo unívocamente. Hace falta un criterio adicional que nos permita elegir uno entre los infinitos operadores que reproducen la información disponible. El criterio que adoptaremos es el siguiente: "elegiremos el operador estadístico que reproduzca toda la verdad y nada más que la verdad". Para concretar este criterio nos valdremos de la función I. Cuanto menor es el valor de I, más información contiene  $\hat{P}$ . Consideremos dos operadores  $\hat{P}_1$  y  $\hat{P}_2$  que satisfagan las condiciones (I.10) y

sean tales que  $I_1 > I_2$ . Entonces  $\hat{P}_2$  contendrá más información que  $\hat{P}_1$ . Pero como ambos reproducen toda la verdad conocida,  $\hat{P}_1$  estará más cerca que  $\hat{P}_2$  de contener nada más que esa verdad. De todos los posibles  $\hat{P}_k$ , el que más cerca esté de satisfacer el criterio que hemos adoptado, será el que tenga asociada máxima ignorancia.

La discusión anterior está contenida en el segundo principio básico:

"Asignaremos a nuestro sistema el operador estadístico que reproduzca toda la información disponible y maximice la falta de información o ignorancia asociada I"

Calcularemos explícitamente el operador basándonos en este segundo principio. Matemáticamente el problema puede definirse como sigue: Se debe hallar el operador  $\hat{P}$  que maximice la expresión:

$$I = -K \text{TR}(\hat{P} \ln \hat{P}) \quad (\text{I.11})$$

sujeta a los vínculos:

$$1) \quad \langle \hat{1} \rangle = \text{TR}(\hat{P}) = 1 \quad , \quad \hat{1} \text{ es el operador unidad} \quad (\text{I.12})$$

$$2) \quad \hat{P} \geq 0 \quad (\text{I.13})$$

$$3) \quad \text{TR}(\hat{P} \hat{O}_i) = \langle \hat{O}_i \rangle = o_i \quad , \quad i=1, \dots, M \quad (\text{I.14})$$

Impondremos las  $M+1$  restricciones dadas por las condiciones 1) y 3) por medio de  $M+1$  multiplicadores de Lagrange, dejando la consideración de la condición 2) para

más adelante.

Para efectuar el proceso variacional con vínculos, variaremos la función:

$$I' = I + K(\lambda_0 + 1) \langle \hat{I} \rangle - K \sum_{i=1}^N \lambda_i \langle \hat{O}_i \rangle \quad (I.15)$$

$$I' = -K \text{TR}[\hat{P}(\ln \hat{P} - (\lambda_0 + 1) + \sum_{i=1}^N \lambda_i \hat{O}_i)] \quad (I.16)$$

La variación hay que efectuarla respecto de  $\hat{P}$ , pero dado que:

$$\hat{P} = \sum_{\alpha} |\alpha\rangle p_{\alpha} \langle \alpha| \quad (I.17)$$

debemos variar tanto respecto de los autoestados  $|\alpha\rangle$  como de los autovalores  $p_{\alpha}$ .

Comenzaremos imponiendo que  $I'$  sea estable frente a un cambio de la base  $|\alpha\rangle$ . Tal cambio es equivalente a una transformación canónica arbitraria de parámetro  $t$  generada por  $\hat{W}$ :

$$\delta \hat{P} = i/\hbar [\hat{W}, \hat{P}] \delta t \quad (I.18)$$

luego:

$$\delta I' = \delta (-K \text{TR}[\hat{P} \ln \hat{P} - \hat{P}(\lambda_0 + 1) + \hat{P} \sum_{i=1}^N \lambda_i \hat{O}_i])$$

$$\delta I' = -K \text{TR}[\delta \hat{P} \ln \hat{P} + \delta \hat{P} - \delta \hat{P}(\lambda_0 + 1) + \delta \hat{P} \sum_{i=1}^N \lambda_i \hat{O}_i] \quad (I.19)$$

$$\delta I' = -K \text{TR}(i/\hbar [\hat{W}, \hat{P}] \ln \hat{P} - i/\hbar [\hat{W}, \hat{P}] \lambda_0 + i/\hbar [\hat{W}, \hat{P}] \sum_{i=1}^N \lambda_i \hat{O}_i) \delta t$$

Los dos primeros términos dentro de la traza son nulos por ser la traza invariante frente a permutaciones cíclicas de los operadores, por lo tanto:

$$\delta I' = -K \text{TR}(i/\hbar [\hat{W}, \hat{P}] \sum_{i=1}^N \lambda_i \hat{O}_i) \delta t \quad (I.20)$$

obviamente esta ecuación se anula si:

$$[\hat{P}, \sum_{i=1}^N \lambda_i \hat{O}_i] = 0$$

y ésta es la única posibilidad de que I sea un extremo para todo  $\hat{W}$ . Esto se ve claramente pues:

$$\text{TR}([\hat{W}, \hat{P}] \sum_{i=1}^N \lambda_i \hat{O}_i) = \text{TR}(\hat{W} [\hat{P}, \sum_{i=1}^N \lambda_i \hat{O}_i])$$

reemplazando en (I.20):

$$\delta I' = -K \text{TR}(\hat{W} i/\hbar [\hat{P}, \sum_{i=1}^N \lambda_i \hat{O}_i]) dt \quad (I.21)$$

Dado que (I.21) debe anularse para todo  $\hat{W}$ , en particular debe hacerlo para:

$$\hat{W} = i/\hbar [\hat{P}, \sum_{i=1}^N \lambda_i \hat{O}_i] \quad (I.22)$$

con esta elección:

$$\delta I' = -K \text{TR}(i/\hbar [\hat{P}, \sum_{i=1}^N \lambda_i \hat{O}_i]^2) \delta t = 0 \quad (I.23)$$

Como la traza del cuadrado de un operador Hermítico es no negativa y sólo se anula cuando el operador es cero,

obtenemos:

$$[\hat{P}, \sum_{i=1}^M \lambda_i \hat{O}_i] = 0 \quad (I.24)$$

como condición para que  $I$  sea estacionaria. Esta relación de conmutación nos dice que  $\hat{P}$  y  $\sum_i \lambda_i \hat{O}_i$  son diagonales en la misma base, o sea que poseen un conjunto común de autoestados, por lo que de acuerdo a (I.17) podemos escribir:

$$\sum_{i=1}^M \lambda_i \hat{O}_i = \sum_{\alpha} |\alpha\rangle \Omega_{\alpha} \langle \alpha| \quad (I.25)$$

Este hecho constituye también condición suficiente para que  $I'$  sea estacionaria respecto de variaciones de la base  $|\alpha\rangle$ . En efecto, sustituyendo (I.17) en (I.16):

$$I' = -K \sum_{\alpha} p_{\alpha} (\ln p_{\alpha} - \lambda_0 - 1 + \langle \alpha | \sum_{i=1}^M \lambda_i \hat{O}_i | \alpha \rangle) \quad (I.26)$$

que es estacionaria respecto de variaciones de los estados  $|\alpha\rangle$  cuando estos son autoestados de  $\sum_i \lambda_i \hat{O}_i$ . Utilizando (I.25):

$$I' = -K \sum_{\alpha} p_{\alpha} (\ln p_{\alpha} - \lambda_0 - 1 + \Omega_{\alpha}) \quad (I.27)$$

Ahora podemos variar respecto de los autovalores  $p_{\alpha}$  sin dificultad:

$$\delta I' / \delta p_{\alpha} = -K (\ln p_{\alpha} - \lambda_0 + \Omega_{\alpha}) = 0 \quad (I.28)$$

de donde:

$$\ln p_{\alpha} = \lambda_0 - \Omega_{\alpha} \quad (I.29)$$

o sea:

$$p_{\alpha} = \exp(\lambda_0 - \Omega_{\alpha}) \quad (I.30)$$

Sustituyendo (I.30) en (I.17) obtenemos el resultado que buscábamos:

$$\hat{P} = \sum_{\alpha} |\alpha\rangle \exp(\lambda_0 - \Omega_{\alpha}) \langle \alpha| = \exp(\lambda_0 - \sum_{i=1}^M \lambda_i \hat{O}_i) \quad (I.31)$$

de la ecuación (I.30) vemos que el vínculo 2)  $\hat{P} \geq 0$  queda satisfecho, pues  $p_{\alpha} > 0$  para todo  $\alpha$  y aunque no lo haremos aquí, es fácil mostrar\* que la expresión de  $\hat{P}$  dada por (I.31), además de hacer estacionaria a I, la maximiza.

Aún quedan por determinar los multiplicadores de Lagrange  $\lambda_i$ ,  $i=0, \dots, M$ . La condición de normalización permite escribir uno de ellos como función de los demás:

$$\begin{aligned} \text{TR}(\hat{P}) &= \text{TR}[\exp(\lambda_0 - \sum_{i=1}^M \lambda_i \hat{O}_i)] \\ &= \exp(\lambda_0) \text{TR}[\exp(-\sum_{i=1}^M \lambda_i \hat{O}_i)] = 1 \end{aligned}$$

de donde:

$$\lambda_0(\lambda_1, \dots, \lambda_M) = -\ln \text{TR}[\exp(-\sum_{i=1}^M \lambda_i \hat{O}_i)] \quad (I.32)$$

o desarrollando la traza:

$$\lambda_0 = -\ln \sum_{\alpha} \exp(-\Omega_{\alpha}) \quad (I.33)$$

El conjunto de vínculos  $\langle \hat{O}_i \rangle = o_i$  puede expresarse como:

$$\delta \lambda_0 / \delta \lambda_i = o_i, \quad i=1, \dots, M \quad (I.34)$$

Este conjunto de ecuaciones junto con la (I.32) permiten determinar los parámetros de Lagrange  $\lambda_i$ ,  $i=0, \dots, M$ , en términos de los datos  $o_i$ .

Es posible demostrar (ver ref. 4) que la solución hallada para  $\hat{F}$  siempre existe y es única si la información de partida no contiene contradicciones. Sin embargo, si los operadores  $\hat{O}_i$  no forman un conjunto linealmente independiente, existirán vínculos adicionales ya que un operador cualquiera del conjunto podrá expresarse como combinación lineal de los demás, y los correspondientes valores medios  $o_i$  deberán satisfacer la misma relación lineal. Si esto no ocurriera, nuestros datos serán contradictorios y no existirá operador estadístico alguno que pueda reproducirlos. De todos modos, aún cuando nuestros datos satisfagan el vínculo lineal que impone  $\{O_i\}$ , la solución para los  $\lambda_i$  no será única ya que podemos agregar en el exponente de  $\hat{F}$  cualquier múltiplo del vínculo lineal (igualado a cero). Sin embargo esto no altera el operador estadístico que está siempre determinado unívocamente si la información que pretendemos reproducir es coherente.

### I.3 RESULTADOS CLASICOS

Desde el punto de vista clásico, el problema del conocimiento del estado de un sistema surge, cuando no se dispone de información suficiente para fijar el punto en el espacio de las fases que lo determina.

En general, tendremos una distribución de probabilidad  $f(q,p)$  y la probabilidad de encontrar el sistema en una región  $R$  del espacio de las fases (EF), será la integral de  $f$  sobre  $R$ .

La función estadística  $f$  refleja nuestro conocimiento sobre el estado del sistema y debe satisfacer:

$$I) \int_{EF} f(q,p) dp_1 dq_1 \dots dp_m dq_m = TR(f) = 1 \quad (I.35)$$

$$II) f(q,p) \geq 0 \quad \text{para todo } p_i, q_i \quad (I.36)$$

Si  $F(q,p)$  es una cantidad dinámica, definimos su promedio o valor de expectación  $\langle F \rangle$  como:

$$\langle F \rangle = TR(f F) \quad (I.37)$$

Las cantidades dinámicas son vectores del espacio vectorial de funciones, mientras que las posibles funciones estadísticas pertenecen al espacio dual del espacio de las cantidades dinámicas. Podemos pensar entonces a la  $TR(f F)$  como un producto escalar:

$$TR(f F) = (f.F) \quad (I.38)$$

Así, la ignorancia  $I$  asociada a la función estadística:



$$I = -K \text{TR}(f \ln f) \quad (I.39)$$

puede escribirse:

$$I = -K(f \cdot \ln f) \quad (I.40)$$

Igual que en el caso cuántico, adoptaremos el segundo postulado para determinar  $f$ , es decir, la elegiremos de manera tal que maximice (I.40) y reproduzca la información disponible, que estará dada en valor medio de cantidades dinámicas  $R_i$ :

$$\langle R_i \rangle = r_i \quad i=1, \dots, M \quad (I.41)$$

La función estadística que cumple con estos requisitos tiene la forma:

$$f(q,p) = \exp(\lambda_0 - \sum_{i=1}^M \lambda_i R_i(q,p)) \quad (I.42)$$

De la condición de normalización se obtiene:

$$\lambda_0 = -\ln \text{TR}(\exp(-\sum_{i=1}^M \lambda_i R_i)) \quad (I.43)$$

Los multiplicadores de Lagrange  $\lambda_i$  que determinan completamente a  $f$ , surgen de resolver el conjunto de ecuaciones (I.41) o sus equivalentes:

$$\delta \lambda_0 / \delta \lambda_i = r_i \quad i=1, \dots, M \quad (I.44)$$

#### I.4 CONCLUSIONES

La Teoría de la Información, da una prescripción bien definida, para construir el operador estadístico  $\hat{P}$  (o la función estadística  $f$  en el caso clásico), partiendo del conocimiento del valor de expectación de  $M$  operadores (o variables dinámicas):

$$\langle \hat{O}_i \rangle = o_i \quad , \quad \langle R_i \rangle = r_i \quad i=1, \dots, M \quad (I.45)$$

y maximizando la falta de información  $I$ :

$$I = -K \text{TR}(\hat{P} \ln \hat{P}) \quad , \quad I = -K \text{TR}(f \ln f) \quad (I.46)$$

El operador (o la función) que maximiza (I.46) sujeto a las restricciones (I.45) está dado por:

$$\hat{P} = \exp\left(\lambda_0 - \sum_{i=1}^M \lambda_i \hat{O}_i\right) \quad (I.47)$$

$$f(q,p) = \exp\left(\lambda_0 - \sum_{i=1}^M \lambda_i R_i(q,p)\right)$$

en término de  $M+1$  multiplicadores de Lagrange  $\lambda_i$ , los cuales son determinados de forma tal que satisfagan las relaciones (I.45) y la condición de normalización:

$$\text{TR}(\hat{P}) = 1 \quad , \quad \text{TR}(f) = 1 \quad (I.48)$$

Esta última condición se utiliza para expresar  $\lambda_0$  en la forma:

$$\lambda_0 = -\ln(\text{TR}(\exp(-\sum_{i=1}^N \lambda_i \hat{O}_i))) \quad ,$$

(I.49)

$$\lambda_0 = -\ln(\text{TR}(\exp(-\sum_{i=1}^N \lambda_i R_i)))$$

Dado un operador  $\hat{O}_k$  (o variable dinámica  $R_k$ ) no incluido en el conjunto (I.45), la teoría predice:

$$\langle \hat{O}_k \rangle = \text{TR}(\hat{\rho} \hat{O}_k) \quad , \quad \langle R_k \rangle = \text{TR}(f R_k) \quad (I.50)$$

De esta forma, podemos evaluar el valor medio de cualquier operador o cantidad dinámica, y estos valores medios serán "nuestras conjeturas".

Fuesto que hemos aceptado los axiomas que definen la TI, supondremos que tales conjeturas son "óptimas", en base a la información disponible.

En realidad, ningún experimento podrá desacreditar nuestra definición de "conjetura óptima", ya que, si las predicciones no son confirmadas empíricamente, argumentaremos que ello es consecuencia de que hemos dejado de lado alguna información relevante. Es decir, los mismos experimentos que nos muestran que nuestras conjeturas están equivocadas, nos proveen nueva información, a partir de la cual podremos predecir. Mientras nos falte alguna información importante, estas nuevas predicciones probablemente también estén equivocadas. A pesar de esto, siempre podremos decir que nada mejor se podría haber hecho.

Aún cuando nuestras conjeturas no sean buenas, podremos afirmar, que son las mejores formulables en base a la

información disponible y esto no podrá refutarse experimentalmente. Esta es la razón por la que adoptamos la definición de "conjetura óptima" construída a partir del segundo principio.

Sin embargo, si realmente se diera el caso, de que nuestras conjeturas fuesen siempre malas, o si para hacer predicciones simples se necesitara información complicada y difícil de obtener, entonces el concepto de "conjetura óptima" tendría poco interés, por más libre que haya sido su elección.

Las evidencias de que disponemos hasta el momento, nos dicen que esto no ocurre. Por el contrario, estamos en condiciones de afirmar que el segundo principio provee una buena forma de conjeturar. En particular, los resultados presentados en este trabajo lo demuestran.

## REFERENCIAS

1. C.E.SHANNON, Bell. Sys. Tech. J.27, 379, 623 (1948)
2. E.T.JAYNES, Phys. Rev. 106, 620 (1957)
3. E.T.JAYNES, Phys. Rev. 108, 171 (1957)
4. A.KATZ, Principles of Statistical Mechanics (Freeman, S.Fco.,1967)
5. A.HOBSON, Concepts in Statistical Mechanics (Gordon and Breach, N.Y., 1971)
6. J.VON NEUMANN, Gott Nachr. 273 (1927)

## CAPITULO I I

### REPRESENTACION DE SEÑALES. METODO DE INVERSION

Si bien una señal contiene información acerca de un sistema observado, en general el procedimiento de detección es el resultado de un proceso y por lo tanto, los datos registrados dependen fuertemente del instrumento de medida utilizado.

Para independizarnos del procedimiento de detección, adoptaremos la representación vectorial de señales<sup>1</sup>, y asociaremos las cantidades observadas a medidas efectuadas sobre las mismas. Esto es, representaremos una señal  $f$  del conjunto de señales de representación positiva como un vector-ket  $|f\rangle$  del espacio de Hilbert y hablaremos de medidas refiriéndonos a aplicaciones de este espacio en el campo de los reales.

En este capítulo, supondremos conocida la base donde las señales admiten la representación que las define, limitándonos a dar una propuesta metodológica, basada en la

TI, que permite obtener dicha representación, a partir del conocimiento parcial que significa disponer de un número finito de medidas.

## II.1 EFECTO DE MUESTREO

Una forma usual de detección de señales, consiste en registrar cantidades que varían como función de algún parámetro, digamos  $x$ . Esta forma de identificar una señal como una función  $f(x)$ , puede expresarse en la representación adoptada, como la proyección de un vector  $|f\rangle$  sobre una dirección particular caracterizada como elemento de una base  $\{|x\rangle\}$  ortogonal y completa del espacio de Hilbert.

La ortogonalidad en el continuo está dada por:

$$\langle x|x'\rangle = \delta(x-x') \quad (\text{II.1})$$

La completitud permite escribir la identidad  $\hat{I}$  de la siguiente manera:

$$\hat{I} = \int |x\rangle \langle x| dx \quad (\text{II.2})$$

Así, una función  $f(x)$  expresada como producto interno  $\langle x|f\rangle$  es un funcional lineal sobre  $|f\rangle$ . En efecto utilizando (II.2):

$$f(x) = \langle x|f\rangle = \int \langle x|x'\rangle \langle x'|f\rangle dx' = \int \delta(x-x') f(x') dx' \quad (\text{II.3})$$

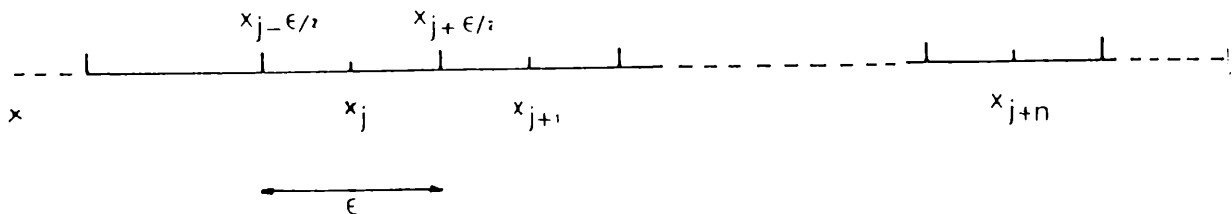
Ahora bien, para poder obtener esta clase de medidas sobre una señal, sería necesario disponer de un aparato "ideal" que sea capaz de proyectar sobre estas direcciones ortogonales. Ningún instrumento real puede satisfacer esta condición, puesto que para hacerlo debería poseer una ventana infinitamente estrecha, tan estrecha como un punto,



sólo así cumplirla con (II.1). Por el contrario, toda ventana posee un determinado ancho, en el que pueden contenerse una cantidad no numerable de puntos, esto hace que sea imposible obtener medidas de una señal como función de un parámetro continuo.

Para visualizar geoméricamente esta afirmación, representaremos las direcciones  $|x\rangle$  como puntos de una recta, cubierta por segmentos de longitud  $\epsilon$ , con los que identificaremos el ancho de una ventana.

Llamaremos  $x_j$  a los puntos que determinan las posiciones de estos segmentos, indicando el punto medio de cada uno de ellos.



Así, una medida efectuada ubicando la ventana en la posición  $x_j$ , contendrá contribuciones provenientes de los puntos pertenecientes al intervalo  $(x_j - \epsilon/2, x_j + \epsilon/2)$ . Si representamos la dirección  $x_j$  como un vector  $|\bar{x}_j\rangle$  éste admitirá una expansión en término de la base  $|x\rangle$ :

$$|\bar{x}_j\rangle = \int_{x_j-\epsilon/2}^{x_j+\epsilon/2} \langle x|\bar{x}_j\rangle |x\rangle dx \quad (\text{II.4})$$

Las cantidades observadas en la posición  $x_j$ , pueden expresarse como el producto interno  $\langle \bar{x}_j|f\rangle$  que se relaciona con  $\langle x|f\rangle$  utilizando (II.4):

$$\langle \bar{x}_j|f\rangle = \int_{x_j-\epsilon/2}^{x_j+\epsilon/2} \langle \bar{x}_j|x\rangle \langle x|f\rangle dx \quad (\text{II.5})$$

El núcleo  $\langle \bar{x}_j|x\rangle$  de esta ecuación, se aleja de la distribución  $\delta(x-x_j)$  en una forma que es característica de cada aparato de medida, pero puede ser siempre representada por una función  $g(x-x_j)$  de norma acotada. Esta propiedad convierte a la ecuación (II.5) en un funcional lineal totalmente continuo<sup>2,3</sup>:

$$\langle \bar{x}_j|f\rangle = \int_{x_j-\epsilon/2}^{x_j+\epsilon/2} g(x-x_j) f(x) dx \quad (\text{II.6})$$

donde:

$$\left( \int_{-\infty}^{\infty} |g(x-x_j)|^2 dx \right)^{1/2} < k \quad \text{para todo } j \quad (\text{II.7})$$

siendo  $k>0$ .

Veamos ahora que los vectores  $|\bar{x}_j\rangle$  forman una base ortogonal, siempre que los segmentos que posicionan la ventana no se superpongan. En efecto:

$$|\bar{x}_k\rangle = \int_{x_k-\epsilon/2}^{x_k+\epsilon/2} \langle x' | \bar{x}_k \rangle |x'\rangle dx' \quad (II.8)$$

$$\langle \bar{x}_j | = \int_{x_k-\epsilon/2}^{x_k+\epsilon/2} \langle \bar{x}_j | x'' \rangle \langle x'' | dx'' \quad (II.9)$$

Por medio de (II.8) y (II.9) podemos expresar  $\langle \bar{x}_j | \bar{x}_k \rangle$  como:

$$\langle \bar{x}_j | \bar{x}_k \rangle = \int_{x_j-\epsilon/2}^{x_j+\epsilon/2} \langle \bar{x}_j | x'' \rangle dx'' \int_{x_k-\epsilon/2}^{x_k+\epsilon/2} \langle x' | \bar{x}_k \rangle \langle x'' | x' \rangle dx' \quad (II.10)$$

pero:

$$\int_{x_k-\epsilon/2}^{x_k+\epsilon/2} \langle x' | \bar{x}_k \rangle \delta(x''-x') dx' = \begin{cases} 0 & \text{si } x'' \notin (x_k-\epsilon/2, x_k+\epsilon/2) \\ \langle x'' | \bar{x}_k \rangle & \text{si } x'' \in (x_k-\epsilon/2, x_k+\epsilon/2) \end{cases} \quad (II.11)$$

por construcción  $\langle \bar{x}_j | = \langle \bar{x}_k |$  si  $x''$  pertenece a  $(x_k-\epsilon/2, x_k+\epsilon/2)$ , luego:

$$\langle \bar{x}_j | \bar{x}_k \rangle = \delta_{jk} \int_{x_k-\epsilon/2}^{x_k+\epsilon/2} |\langle \bar{x}_k | x'' \rangle|^2 dx'' \quad (II.12)$$

Este resultado no es sorprendente, sólo refleja el hecho de que los vectores  $|\bar{x}_j\rangle$  fueron construidos a partir de subconjuntos ortogonales entre sí. No ocurriría lo mismo si hubiéramos permitido que dos segmentos se superpongan, lo que significaría elegir un paso de corrimiento de ventana inferior a su propio ancho.

Tenemos ahora una base  $\{|\bar{x}_j\rangle\}$  ortogonal y discreta, con la que podemos representar las medidas  $f_j$ , efectuadas por un aparato particular, caracterizado por su ancho instrumental  $g(x-x_j)$ :

$$f_j = \langle \bar{x}_j | f \rangle = \int_{x_j - \epsilon/2}^{x_j + \epsilon/2} g(x-x_j) f(x) dx \quad (\text{II.13})$$

Supongamos que disponemos de  $M$  medidas  $f_j$ , (con  $j=1, \dots, M$ ) y lo que pretendemos es obtener la representación de la señal como función de un parámetro que sea expresable por un punto del espacio de Hilbert. En este caso lo que debemos hacer es "deconvolucionar" la señal. Es decir, invertir el sistema de  $M$  ecuaciones del tipo (II.13) construido a partir de  $M$  datos experimentales.

Cualquier intento de inversión presupone discretizar la integral, convirtiéndola en:

$$f_j = \sum_k g(x_k - x_j) f(x_k) \Delta x_k \quad (\text{II.14})$$

Pero si conociéramos  $g(x-x_j)$  como función continua, podríamos hacer  $\Delta x_k$  tan pequeño como deseáramos. En ese caso tendríamos un sistema donde el número de incógnitas  $f(x_k)$  es mayor que el número  $M$  de ecuaciones y la inversión lineal admitiría más de una solución.

Queda por recurrir entonces, a los métodos paramétricos, pero para ello es necesario conocer o suponer una forma funcional de la solución y ajustar los parámetros que la

determinan a partir de las medidas experimentales. En muchos casos, no se conoce con certeza dicha forma funcional y por lo tanto se recurre a modelos o a aproximaciones que permiten "suponerla". Pero en casos todavía más frecuentes, ni siquiera es posible realizar estas suposiciones, por no poseer ninguna evidencia sobre una posible forma funcional. La propuesta de este trabajo, para encontrar una solución al problema planteado, es adoptar el criterio que provee la TI y elegir una forma funcional que nos garantice (dentro de los postulados de la teoría) la solución menos prejuiciosa, es decir, la que reproduzca la información de que disponemos, y nada más que eso.

En la sección IV introduciremos la forma de instrumentar el criterio adoptado.

## II.2 SUPERPOSICION DE SEÑALES

Puesto que una señal refleja de alguna manera el estado de un sistema físico, cuando el sistema está definido como superposición de otros estados, la señal que se le asocia tendrá que poder ser expresada como superposición de las señales que producirían independientemente cada uno de estos estados.

Llamaremos  $\{|r_n\rangle\}$  al conjunto de dichas señales, identificándolas con un conjunto de vectores a los que le exigiremos independencia lineal. De esta manera, una señal  $|f\rangle$  que es respuesta del sistema mencionado podrá expresarse:

$$|f\rangle = \sum_n C_n |r_n\rangle \quad (\text{II.15})$$

Los coeficientes  $C_n$  de la expansión anterior, representan el peso de cada señal  $|r_n\rangle$  y por lo tanto sólo pueden ser cantidades reales y positivas.

A pesar de que a la base  $\{|r_n\rangle\}$  no le exigimos ortogonalidad, siempre es posible encontrar otra base que sea bi-ortogonal a ésta. La bi-ortogonalidad está dada en un sentido tal que los vectores  $|w_n\rangle$  que la constituyen cumplen con la condición<sup>4</sup>:

$$\langle w_n | r_m \rangle = \delta_{nm}. \quad (\text{II.16})$$

La existencia de esta base bi-ortogonal está garantizada por la independencia lineal de los vectores  $|r_n\rangle$ . En

efecto, si desarrollamos estos vectores en términos de la nueva base podemos encontrar la matriz de transformación:

$$|r_n\rangle = \sum_k \langle r_k | r_n \rangle |w_k\rangle$$

La inversión de este sistema podrá hacerse, puesto que la matriz de los productos internos  $\langle r_k | r_n \rangle$  es no singular como consecuencia de la independencia lineal de los vectores. Por lo tanto, en el caso de dimensión finita siempre podremos encontrar la base en la que proyectar  $|f\rangle$  para obtener los coeficientes  $C_n$  que dan el peso de cada señal. Esto nos permite reescribir (II.15) en la forma:

$$|f\rangle = \sum_n \langle w_n | f \rangle |r_n\rangle \quad (\text{II.17})$$

La completitud de la base permite expresar la identidad en ese subespacio con:

$$\hat{I} = \sum_n |r_n\rangle \langle w_n| \quad (\text{II.18})$$

Si de todas las posibles medidas sobre  $|f\rangle$  nos quedamos sólo con aquellas que pueden expresarse como funcionales lineales  $\{l_i\}$ , dichas medidas, además de aplicar el espacio de señal en números, satisfarán la condición:

$$l_i(|f\rangle) = \sum_n \langle w_n | f \rangle l_i(|r_n\rangle) \quad (\text{II.19})$$

Introducir (II.16) en (II.19) nos permite pasar a otra representación más conveniente:

$$I_1(|f\rangle) = \sum_n \langle W_n | f \rangle \langle W_n | I_1(|r_n\rangle) |r_n\rangle \quad (II.20)$$

Esta expresión puede considerarse como la traza del producto de dos operadores ( $\hat{P}'$ ,  $\hat{L}_1$ ) cada uno de los cuales está definido de la siguiente manera:

$$\hat{P}' = \sum_n |r_n\rangle \langle W_n | f \rangle \langle W_n | \quad (II.21)$$

$$\hat{L}_1 = \sum_n |r_n\rangle I_1(|r_n\rangle) \langle W_n | \quad (II.22)$$

de modo que:

$$\begin{aligned} \text{TR}(\hat{P}' \hat{L}_1) &= \sum_n \langle W_n | \hat{P}' \hat{L}_1 |r_n\rangle \\ &= \sum_{nn'} \langle W_n | r_{n'} \rangle \langle W_n' | f \rangle \\ &\quad \langle W_n' | r_{n''} \rangle I_1(|r_{n''}\rangle) \langle W_{n''} | r_n \rangle \\ &= \sum_{nn'} \delta_{nn'} \langle W_n' | f \rangle \\ &\quad \delta_{n'n''} I_1(|r_{n''}\rangle) \delta_{n''n} \\ &= \sum_n \langle W_n | f \rangle I_1(|r_n\rangle) \\ &= I_1(|f\rangle) \end{aligned} \quad (II.23)$$

En el caso en que los vectores  $|r_n\rangle$  formen base ortonormal, la representación de la señal, dada por el



conjunto de coeficientes  $C_n$ , será igual a las proyecciones sobre la base de señales, de lo contrario estará dada por la proyección sobre la base bi-ortogonal a ella, también denominada recíproca.

Las definiciones (II.21) y (II.22) responden a lo dicho en el párrafo anterior, puesto que si los vectores  $|r_n\rangle$  son ortonormales, la base que constituyen es autorrecíproca, es decir, coincide con su recíproca. La ecuación (II.23) nos permite expresar ambos casos de manera general y será útil en el desarrollo del método para encontrar la representación de la señal.

### II.3 IGNORANCIA ASOCIADA A LA REPRESENTACION DE SEÑALES

A pesar de que en el conjunto de números  $\{l_i(|r_n\rangle)\}$  están contenidos los resultados posibles de una medida particular  $l_i$  sobre las señales  $|r_n\rangle$ , la ecuación (II.19), o su equivalente (II.23), refleja el hecho de que sólo es posible conocer parcialmente el mecanismo que determina el valor de esta medida.

Si a cada vector  $|r_n\rangle$  le asociamos un evento elemental, al funcional lineal  $l_i$  que transforma estos eventos en números reales, podemos considerarlo una variable aleatoria que puede tomar los valores  $\{l_i(|r_n\rangle)\}$  de acuerdo a una distribución caracterizada por el conjunto de coeficientes  $\{C_n = \langle w_n | f \rangle\}$ . A pesar de que este conjunto de coeficientes no está normalizado, siempre es posible, a partir de él, encontrar otro conjunto normalizado. Es decir, podemos escribir cada  $C_n$  como:

$$C_n = A p_n \quad (II.24)$$

Donde  $A$  es una constante positiva y los  $p_n$  están normalizados.

Esta nueva distribución satisface todas las condiciones requeridas por una distribución de probabilidades:

$$p_n \geq 0 \quad \text{para todo } n \quad (II.25)$$

$$\sum_n p_n = 1$$

Esto nos permite asignar a los posibles valores de una medida  $l_i$  las correspondientes probabilidades  $\{p_n\}$ .

Así, asociando los vectores con eventos, hemos convertido el espacio vectorial de señales en un espacio muestral, que designaremos  $S$ . A una medida particular  $l_i$  la consideraremos variable aleatoria de este espacio  $S$  en el conjunto  $l_i(S) = \{l_i(|r_n\rangle)\}$  y luego convertiremos  $l_i(S)$  en un espacio de probabilidad, definiendo la probabilidad de  $l_i(|r_n\rangle)$  como:

$$p_n = \langle w_n | f \rangle / \sum_n \langle w_n | f \rangle \quad (\text{II.26})$$

Desde este punto de vista, el resultado de una medida sobre una señal  $|f\rangle$ , es proporcional al valor medio de dicha medida, considerada como variable aleatoria:

$$l_i(|f\rangle) \sim \sum_n p_n l_i(|r_n\rangle) = \langle l_i \rangle \quad (\text{II.27})$$

más precisamente:

$$l_i(|f\rangle) = A \langle l_i \rangle \quad (\text{II.28})$$

o:

$$l_i(|f\rangle) = A \text{TR}(\hat{P} \hat{L}_i) = A \langle \hat{L}_i \rangle \quad (\text{II.29})$$

donde :

$$\hat{P} = \hat{P}' / \text{TR}(\hat{P}') = \sum_n |r_n\rangle p_n \langle w_n| \quad (\text{II.30})$$

Si contáramos con un dispositivo capaz de efectuar las medidas  $\{l_i(|r_n\rangle)\}$ , sólo podríamos asegurar el resultado

de una medida particular cuando una sola de las proyecciones  $\langle w_n | f \rangle$  fuera distinta de cero. En cualquier otro caso, tendríamos que hablar de la probabilidad de obtener un resultado particular. Esto reflejaría cierta ignorancia sobre el sistema. Adoptaremos como expresión de esta ignorancia a la función I definida en el capítulo I:

$$I = -K \sum_n p_n \ln p_n \quad (\text{II.31})$$

Por estricta analogía con la mecánica estadística, llamaremos a  $\hat{P}$  operador estadístico y reescribiremos (II.31) como:

$$I = -K \text{TR}(\hat{P} \ln \hat{P}) \quad (\text{II.32})$$

## II.4 METODO DE INVERSION

Encontrar la representación de una señal a partir de un número finito de medidas sobre la misma, digamos  $M$ , presupone resolver un conjunto de ecuaciones de la forma:

$$f_i = A \text{TR}(\hat{F} \hat{L}_i) \quad \text{con } i=1, \dots, M \quad (\text{II.33})$$

donde:

$$f_i = 1_i \langle 1f \rangle$$

Puesto que  $A$  es una constante en general no conocida, es necesario utilizar una ecuación para despejarla. Tomando por ejemplo la primera:

$$A = f_1 / \text{TR}(\hat{F} \hat{L}_1) \quad (\text{II.34})$$

podemos reescribir el sistema (II.33) de la siguiente manera:

$$\text{TR}(\hat{F} \hat{O}_i) = \langle \hat{O}_i \rangle = 0 \quad \text{con } i=2, \dots, M \quad (\text{II.35})$$

donde los operadores  $\hat{O}_i$  están definidos como:

$$\hat{O}_i = f_i \hat{L}_1 - f_1 \hat{L}_i \quad \text{con } i=2, \dots, M \quad (\text{II.36})$$

Para encontrar el operador estadístico  $\hat{F}$ , adoptaremos los postulados de la TI discutidos en el capítulo I. Siguiendo la prescripción que da la teoría, el operador estadístico que hace máxima la falta de información (II.32) sujeto a las restricciones (II.35) tiene la forma:

$$\hat{P} = \exp(\lambda_0 - \sum_{i=2}^M \lambda_i \hat{O}_i) \quad (\text{II.37})$$

Agregando la condición de normalización:

$$\text{TR}(\hat{P}) = 1 \quad (\text{II.38})$$

podemos despejar  $\lambda_0$ :

$$\lambda_0 = -\ln \text{TR}(-\exp(-\sum_{i=2}^M \lambda_i \hat{O}_i)) \quad (\text{II.39})$$

Los multiplicadores de Lagrange que determinan el operador  $\hat{P}$ , surgen de resolver el conjunto de ecuaciones (II.35) o sus equivalentes:

$$\delta \lambda_0 / \delta \lambda_i = 0 \quad i=2, \dots, M \quad (\text{II.40})$$

Una vez encontrado este operador, a partir de un número  $M$  de medidas, la teoría nos permite predecir el valor de otras medidas no utilizadas para construir las ecuaciones (II.35). Por ejemplo, podemos predecir las medidas  $M+j$  por medio de:

$$f_{M+j}^n = f_1 \text{TR}(\hat{P} \hat{L}_{M+j}) / \text{TR}(\hat{P} \hat{L}_1) \quad (\text{II.41})$$

El superíndice  $M$  colocado a  $f$ , significa que a la medida  $M+j$  la predecimos tomando  $M$  valores experimentales como información de partida. Si las medidas predichas coinciden con las experimentales en el grado de aproximación deseado, estaremos en condiciones de afirmar que las medidas utilizadas para construir el operador estadístico, dan

suficiente información sobre el sistema, puesto que permiten predecir correctamente. Por el contrario, obtener por medio de (II.41) medidas que no coinciden con las experimentales, significará que la información de partida no ha sido suficiente como para dar una descripción correcta. En este caso debemos aumentar el número de ecuaciones (II.35) hasta que logremos las predicciones deseadas.

## II.5 CONVERGENCIA DENTRO DE LA INFORMACION DISPONIBLE

Supongamos disponer de un número  $k$  de medidas experimentales  $\{f_i\}$ . Tomando sólo dos de estas medidas construimos el operador  $\hat{F}^1$ , luego agregamos una nueva medida para construir  $\hat{F}^2$ . Procediendo de la misma manera, generamos la sucesión  $\{\hat{F}^n\}$ . A partir de esa sucesión de operadores obtenemos  $k$  sucesiones de medidas por medio de:

$$f_i^M = f_i \text{ TR}(\hat{F}^M \hat{L}_i) / \text{TR}(\hat{F}^M \hat{L}_i) \quad i=1, \dots, k \quad (\text{II.42})$$

Diremos que el operador estadístico converge a la representación de la señal, dentro de la información disponible, si toda sucesión de medidas construídas a partir de él, converge a las medidas experimentales dentro del error con que éstas se conocen. Es decir, si dado  $\epsilon_i > 0$ , existe  $M_0$  tal que:

$$|f_i - f_i^M| < \epsilon_i \quad \text{para todo } M \geq M_0, \text{ con } i=1, \dots, k \quad (\text{II.43})$$

Podemos asegurar este tipo de convergencia si admitimos el problema bien planteado (es decir, si la señal  $|f\rangle$  pertenece a la imagen de la transformación (II.15)) y si no existe dependencia lineal entre las medidas  $l_i$ . Recordemos que cada  $l_i$  es un funcional lineal particular, y el conjunto de funcionales lineales sobre un espacio vectorial forma, en si mismo, un espacio vectorial. Por lo tanto, si ningún funcional lineal pertenece al subespacio generado por el resto, estaremos seguros que la información no es



contradictoria, lo que garantiza la existencia y unicidad del operador estadístico<sup>8</sup>. Luego, al menos para  $M=M_0=k$  se verificará (II.43). Sin embargo, los resultados que presentamos en los capítulos III y V mostrarán que si el número  $k$  de medidas disponibles no es extremadamente pequeño, la convergencia se logra para  $M_0 \ll k$ .

## II.6 CONCLUSIONES

Partiendo del conocimiento que se tiene "a priori" sobre el signo definido de la representación de una señal, hemos adoptado la prescripción que da la TI para asumir que dicha representación es de la forma:

$$C_n = f_1 \exp\left(-\sum_{j=2}^M \lambda_j O_j(n)\right) / \left\{ \sum_n \exp\left(-\sum_{j=2}^M \lambda_j O_j(n)\right) L_1(n) \right\} \quad (\text{II.44})$$

Los multiplicadores que determinan esta representación, surgen de exigirle a la misma que reproduzca  $M$  medidas experimentales  $f_i$  (con  $i=1, \dots, M$ ). Esto se logra resolviendo el conjunto de ecuaciones:

$$\sum_n \exp\left(-\sum_{j=2}^M \lambda_j O_j(n)\right) O_i(n) = 0 \quad i=2, \dots, M \quad (\text{II.45})$$

donde:

$$O_i(n) = f_i L_1(n) - f_1 L_i(n) \quad (\text{II.46})$$

y

$$L_i(n) = l_i(r_n)$$

Conocidos  $M$  multiplicadores de Lagrange, podremos "conjeturar" otras medidas por medio de la siguiente ecuación:

$$f_{j,m} = f_1 \sum_n \exp\left(-\sum_{i=2}^M \lambda_i O_i(n)\right) L_{j,m}(n) / \left\{ \sum_n \exp\left(-\sum_{i=2}^M \lambda_i O_i(n)\right) L_1(n) \right\} \quad (j=1, \dots, \infty) \quad (\text{II.47})$$

Si disponemos de  $k$  medidas experimentales, podemos compararlas con nuestras predicciones tomando  $j=1, \dots, (k-M)$  en (II.47). Si estos valores coinciden, dentro del error experimental, diremos que (II.44) converge a la representación de la señal, dentro de la información disponible. Si la diferencia entre las medidas conocidas y las conjeturadas es mayor que el error experimental, tendremos que incrementar el número  $M$  de ecuaciones (II.45) hasta lograr la convergencia deseada.

Es claro que el tipo de convergencia que hemos definido, no significa convergencia a la solución del problema. Por ejemplo, si la información disponible es escasa, de modo que sólo se satisface (II.43) para  $M_0=k$ , seguramente las predicciones que hagamos a partir de  $\hat{f}^*$  serán incorrectas (aunque no tengamos forma de verificarlo). Sin embargo, puesto que hemos adoptado el criterio de "conjetura óptima" dado por el segundo principio, podemos asegurar que nada mejor podría haberse hecho a partir de esa información disponible. Además, ninguna nueva medida podrá contradecir esta afirmación. Una nueva medida sólo aumenta la información que tenemos y si no confirma nuestras predicciones deberá incorporarse como nueva restricción.

Como ya lo hemos discutido anteriormente, a pesar de que ninguna medida puede desacreditar la definición subjetiva de "conjetura óptima", si para poder predecir bien necesitáramos partir siempre de una gran cantidad de información (lo que implica resolver un número grande de

ecuaciones no lineales), entonces el método propuesto sería impráctico. En cambio, si a partir de poca información podemos predecir correctamente, es decir si la convergencia se alcanza para  $M_0 \ll k$ , entonces el método será útil y aún en los casos en que no dispongamos de datos que verifiquen nuestras predicciones, tendremos derecho a decir que la solución obtenida es la mejor que podemos formular en base a la información de que disponemos. Dado que la única manera de evaluar nuestra forma de conjeturar es confrontándola con datos conocidos, en el siguiente capítulo resolveremos algunos ejemplos que permiten efectuar estas comparaciones.

## REFERENCIAS

1. C.E. SHANNON, Communication in the Presence of Noise  
Proc. IRE, Vol. 37, N° 1
2. A.N. KOLMOGOROV, S.V. FOMIN, Elementos de la Teoría de  
funciones y del análisis funcional MIR (1975)
3. M. REED, B. SIMON, Methods of Modern Mathematical  
Physics. I: Functional Analysis. Academic Press, Inc.  
(1980)
4. P.M. MORSE, H. FESHBACH, Methods of Theoretical Physics,  
Vol I. (Mc. Graw-Hill, New York, 1953)
5. A. KATZ, Principles of Statistical Mechanics, New York:  
Freeman (1967)

## **CAPITULO III**

### **ALGUNAS APLICACIONES**

Este capítulo lo dedicaremos a dar algunas aplicaciones del método de inversión presentado en el capítulo II. A efectos de mostrar la rápida convergencia a la representación buscada, resolveremos algunos ejemplos numéricos cuya solución es conocida. Aplicaremos el método considerando dos clases de medidas lineales. Una de ellas supone el conocimiento de medidas puntuales, y la otra, el conocimiento de medidas que se asocian con los momentos de una distribución.

### III.1 INVERSION A PARTIR DE MEDIDAS PUNTUALES'

Aceptando la prescripción dada por la TI, hemos asumido que el operador estadístico que nos permite encontrar la representación positiva de una señal es de la forma:

$$\hat{F} = \exp(\lambda_0) \exp\left(-\sum_{i=2}^M \lambda_i \hat{O}_i\right) \quad (\text{III.1})$$

Para definir los operadores  $\hat{O}_i = f_i \hat{L}_i - f_i \hat{L}_i$ , hemos supuesto que las medidas  $f_i$  sobre  $|f\rangle$  son representables por medio de funcionales lineales,  $f_i = l_i(|f\rangle)$ .

En esta sección particularizaremos cada funcional lineal  $l_i$ , aceptando que una medida  $f_i$  es un valor particular obtenido proyectando  $|f\rangle$  sobre la dirección  $\langle x_i|$ , es decir:

$$f_i = \langle x_i | f \rangle \quad (\text{III.2})$$

y

$$\hat{L}_i = \sum_n |r_n\rangle \langle x_i | r_n \rangle \langle w_n | \quad (\text{III.3})$$

Tomando  $M$  medidas de la forma (III.2), la representación  $C_n$  de la señal, estará dada por:

$$C_n = \langle x_1 | f \rangle \exp\left(-\sum_{i=2}^M \lambda_i O_i(n)\right) / \left\{ \sum_n \exp\left(-\sum_{i=2}^M \lambda_i O_i(n)\right) \langle x_1 | r_n \rangle \right\} \quad (\text{III.4})$$

con:

$$O_1(n) = \langle x_1 | f \rangle \langle x_1 | r_n \rangle - \langle x_1 | f \rangle \langle x_1 | r_n \rangle \quad (\text{III.5})$$

Los ejemplos de inversión en que intervienen esta clase de medidas puntuales serán desarrollados en las secciones (III.3) y (III.5).



### III.2 INVERSION A PARTIR DE LOS MOMENTOS DE UNA DISTRIBUCION\*

No siempre es posible disponer de medidas puntuales sobre una señal. Existen dispositivos de proceso que miden intensidades integradas, el valor máximo de un pico, su desviación cuadrática media, etc. Dado que estas medidas pueden relacionarse con los momentos de una distribución, proponemos aplicar el método de inversión suponiendo conocidos dichos momentos.

En este caso, tomando  $M+1$  medidas de la forma:

$$(x-x_0)^k \langle x|f \rangle dx = m_k \quad \text{con } k=0, \dots, M \quad (\text{III.6})$$

podemos obtener la representación de la señal, por medio de la expresión:

$$C_n = m_0 \exp\left(-\sum_{k=1}^M \lambda_k O_k(n)\right) / \left\langle \sum_n \exp\left(-\sum_{k=1}^M \lambda_k O_k(n)\right) M_0(n) \right\rangle \quad (\text{III.7})$$

Puesto que hemos utilizado el momento de orden cero para despejar la constante  $A$ , los  $O_k(n)$  resultan:

$$O_k(n) = m_k M_0(n) - m_0 M_k(n) \quad (\text{III.8})$$

con:

$$M_k(n) = \int (x - x_0)^k \langle x|r_n \rangle dx \quad (\text{III.9})$$

### III.3 EJEMPLO NUMERICO

Para comparar nuestra propuesta con otros métodos, resolveremos la siguiente inversión simulada numéricamente:

$$f_i = \int_a^b K_i(u) f(u) du \quad (\text{III.10})$$

donde la representación buscada es:

$$f(u) = 1 + 4(u - 0.5)^2 \quad (\text{III.11})$$

y

$$K_i(u) = u \exp(-y_i u) \quad (\text{III.12})$$

En la ecuación (III.10) se tomó  $a = 0$  y  $b = 1$  y se discretizó la integral con un  $\Delta u = 0.01$ . Así:

$$f_i = \sum_n K_i(u_n) \bar{f}(u_n) \Delta u_n \quad (\text{III.13})$$

donde:

$$u_n = (n-1) \Delta u$$

Considerando que  $\bar{f}(u_n)$  es el valor medio de  $f(u_n)$  en el  $n$ -ésimo intervalo y  $C_n = \bar{f}(u_n) \Delta u_n$ , reescribiremos (III.13):

$$f_i = \sum_n C_n K_i(u_n) \quad (\text{III.14})$$

En las figuras III.1.a, III.1.b y III.1.c se muestran los

resultados logrados por tres métodos diferentes. Estos resultados están publicados en ref. 3. Los métodos empleados son los siguientes:

**Fig. III.1.a Inversión lineal restringida<sup>4,6</sup>**

**Fig. III.1.b Método de síntesis<sup>4,7</sup>**

**Fig. III.1.c Iteración no lineal<sup>8,9</sup>**

Las figuras III.1.d y III.1.e muestran los resultados obtenidos por el método propuesto en este trabajo. En la figura III.1.d podemos ver la solución  $f^3(u_n)$  obtenida a partir de cuatro puntos:  $y_1=0$ ,  $y_2=1$ ,  $y_3=5$ ,  $y_4=10$ . La expresión para  $f^3(u_n)$ :

$$\begin{aligned}
 f^3(u_n) = & f_1 \exp(-\lambda_1 O_1(u_n) - \lambda_2 O_2(u_n) - \\
 & - \lambda_3 O_3(u_n)) / \left\{ \sum_n \exp(-\lambda_1 O_1(u_n) - \right. \\
 & \left. - \lambda_2 O_2(u_n) - \lambda_3 O_3(u_n)) K_1(u_n) \right\} \quad (III.15)
 \end{aligned}$$

donde:

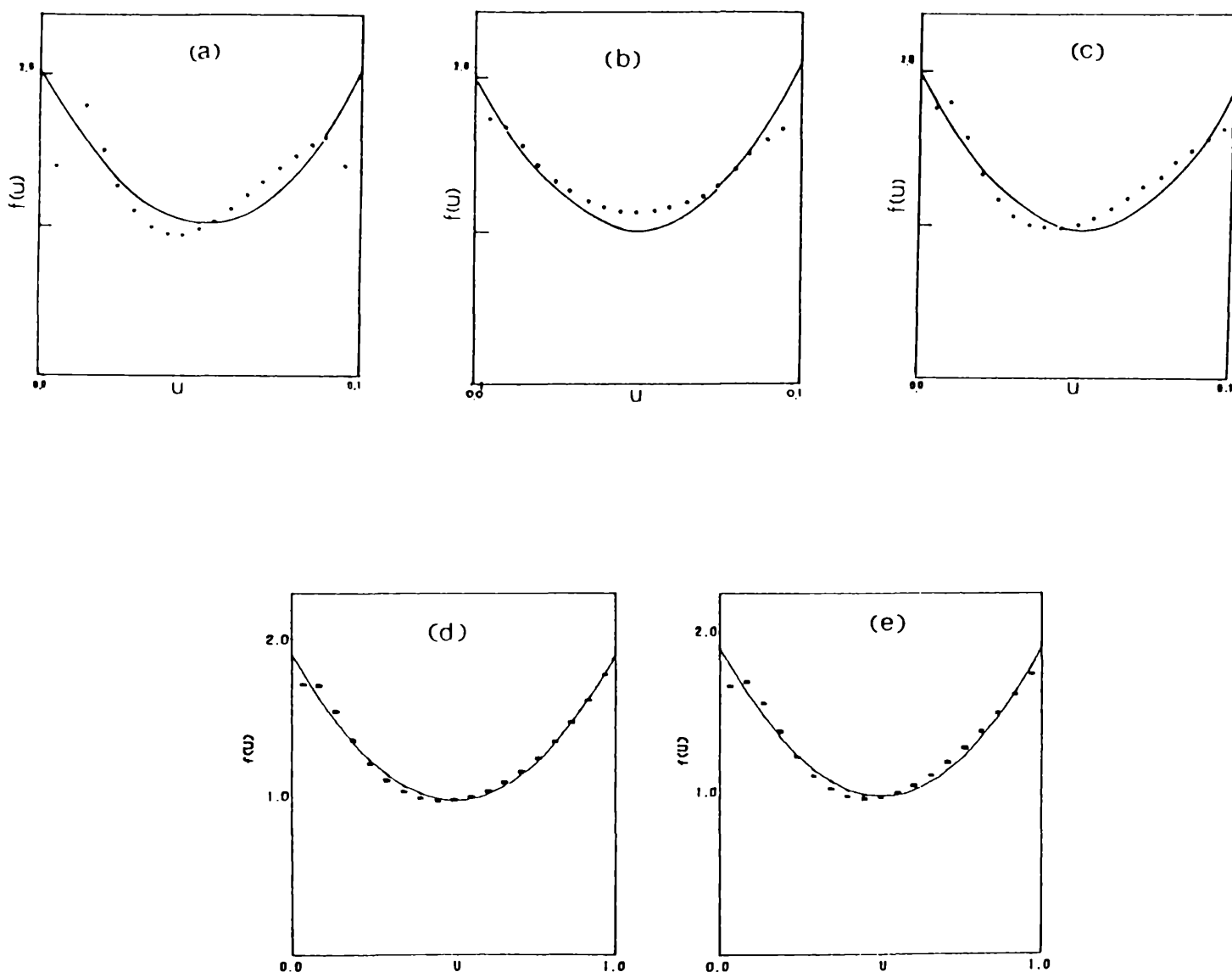
$$\begin{aligned}
 O_1(u_n) = & f_2 K_1(u_n) - K_2(u_n) f_1 \\
 = & f_2 u_n - f_1 \exp(-u_n) u_n
 \end{aligned}$$

$$\begin{aligned}
 O_2(u_n) = & f_3 K_1(u_n) - f_1 K_3(u_n) \\
 = & f_3 u_n - f_1 \exp(-5 u_n) u_n
 \end{aligned}$$

$$\begin{aligned}
 O_3(u_n) = & f_4 K_1(u_n) - f_1 K_4(u_n) \\
 = & f_4 u_n - f_1 \exp(-10 u_n) u_n
 \end{aligned}$$

Los valores de los multiplicadores que determinan (III.15) son  $\lambda_1 = -7.2449$ ,  $\lambda_2 = 1.7905$ ,  $\lambda_3 = 16.1179$ .

La figura III.1.e muestra la solución obtenida a partir de los momentos de orden 0, 2 y 4, calculados en el intervalo  $0 < \gamma_i < 6.5$ . Los multiplicadores que resultaron son  $\lambda_1 = -1.0544$ ,  $\lambda_2 = 0.040$ .



**Fig. III.1:** Inversión simulada de  $f(u) = 1 - 4(u - 0.5)^2$ . a) Inversión lineal. b) Método de síntesis. c) Iteración no lineal. d) solución obtenida por TI a partir de cuatro medidas puntuales ( $\gamma_1 = 0$ ,  $\gamma_2 = 1$ ,  $\gamma_3 = 5$ ,  $\gamma_4 = 10$ ). e) Solución por TI a partir de tres momentos calculados en el intervalo  $(0, 6.5)$ .  $\mu_0 = 1.09$ ,  $\mu_2 = 5.75$ ,  $\mu_4 = 108$ .

### III.4 DECONVOLUCION DEL ANCHO INSTRUMENTAL<sup>2</sup>

#### III.4.1 SIMULACION

Antes de tratar un caso realista, resolveremos un ejemplo numérico, considerando que la función  $g(x)$  de la ecuación (II.14) tiene la siguiente expresión analítica:

$$g(x) = 1/(1+(K_1x)^2) + 0.5/(1+(K_2(x-0.4))^2)$$

donde:

$$K_1 = 2/0.18 \quad \text{y} \quad K_2 = 2/0.3$$

Para encontrar la representación  $f(x)$  buscada se tomaron los momentos de orden cero y dos. Como se ve en **Tabla III.1**, la información que proveen estos momentos no es suficiente, ya que las predicciones  $f_i^1$  que se logran no concuerdan satisfactoriamente con los valores conocidos. Puesto que los momentos impares tampoco son bien reproducidos, se agregó el momento de primer orden como información de partida. Con esta nueva representación, determinada por los multiplicadores  $\lambda_1 = -0.1664$  y  $\lambda_2 = 0.944$ , se alcanzó una muy buena concordancia entre los valores predichos  $f_i^2$  y los conocidos  $f_i$ .

TABLE III.1

x	$f_i$	$f_i^1$	$f_i^2$
-1.0	6.59	11.08	6.69
-0.9	8.09	16.89	8.09
-0.8	10.26	25.74	10.26
-0.7	13.65	40.68	13.65
-0.6	19.57	63.96	19.55
-0.5	30.93	96.12	30.89
-0.4	53.06	134.27	53.00
-0.3	91.85	171.76	91.85
-0.2	146.76	200.51	146.90
-0.1	203.39	214.81	203.62
0.0	239.24	213.72	239.34
0.1	243.38	200.11	243.26
0.2	225.96	177.98	225.81
0.3	202.83	150.69	202.82
0.4	178.00	121.05	178.05
0.5	147.19	91.99	147.18
0.6	110.84	63.35	110.87
0.7	76.07	46.00	76.02
0.8	49.30	31.33	49.30
0.9	31.83	21.52	31.85
1.0	21.41	15.25	21.43

### III.4.2 CASO REALISTA: DECONVOLUCION DE UN PERFIL PRODUCIDO POR DIFRACCION DE RAYOS X.

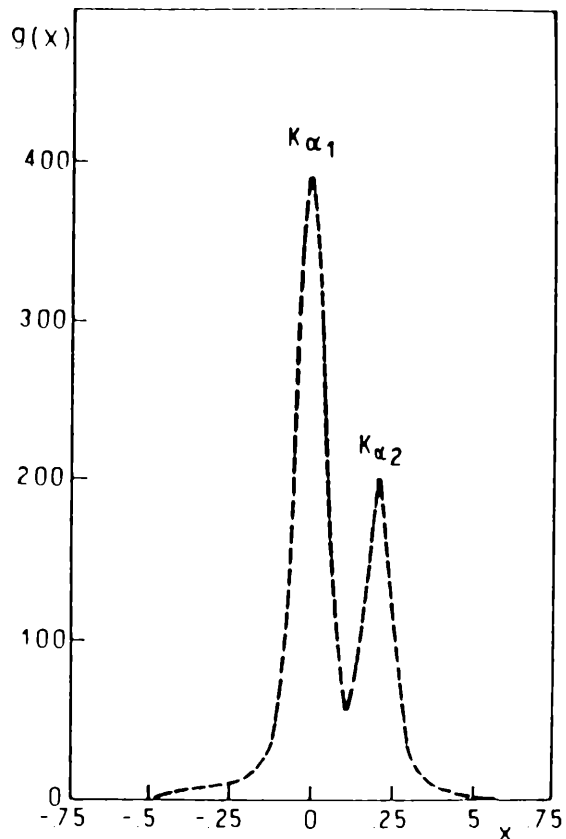
Recordemos que en la ecuación (II.14):

$$f_1 = \sum_k g(x_1 - x_k) f(x_k) \Delta x_k$$

Los valores  $f(x_k)$  representan las medidas que únicamente podrían efectuarse con un aparato ideal, cuyo ancho instrumental fuese una delta. A pesar de que bajo algunas condiciones particulares, despreciar el ancho instrumental no conduce a errores significativos, en otros casos es imposible despreciarlo.

En las experiencias de difracción de Rayos X, se incluyen en la función  $g(x-x_k)$ , todos los efectos que producen distanciamiento de las condiciones consideradas ideales. De la ecuación (II.13) se deduce que esta función podría ser medida, si contáramos con una muestra ideal de la que se espera un perfil de difracción tendiente a una delta.

Para evaluar cualitativamente los resultados de la deconvolución aplicada a un caso realista, se tomaron para el ancho instrumental las medidas realizadas sobre una lámina de cobre calentada a 400 °C. En estas condiciones podemos considerar que el ancho de línea se debe solamente a efectos instrumentales. Su espectro compuesto de 50 puntos equiespaciados en  $\Delta 2\theta = 0.02$ , se ha graficado en la figura III.2, donde puede observarse la separación de las líneas  $K_{\alpha 1}$ ,  $K_{\alpha 2}$ .



**Fig. III.2:** Ancho instrumental donde se puede observar claramente el doblete  $K\alpha_1$ ,  $K\alpha_2$

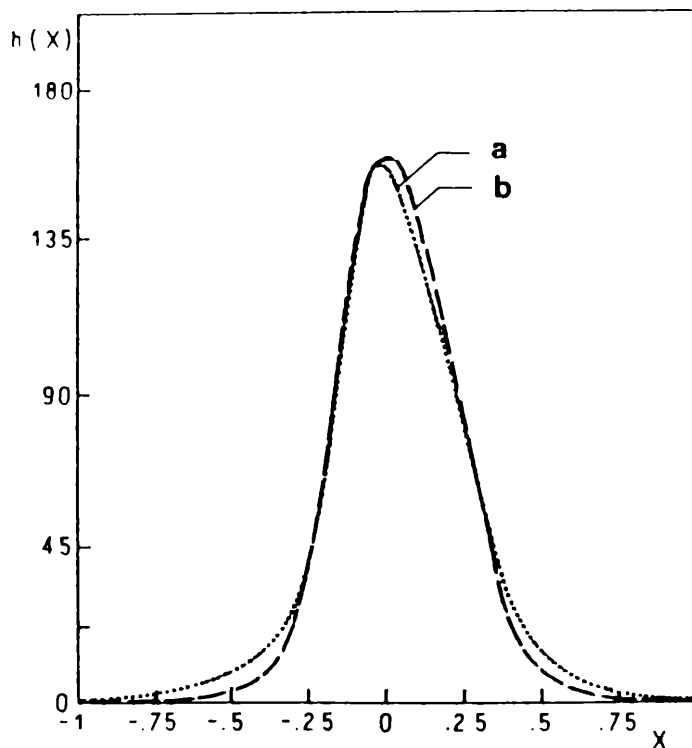
En el ejemplo experimental elegido\*, el perfil  $f_i$  graficado en la figura III.3.a fue tomado de la reflexión correspondiente a los planos (220) de otra lámina de cobre.

\* Las medidas fueron realizadas por el grupo de Física del Estado Sólido, del Centro de Investigación y Desarrollo en Procesos Catalíticos (CINDECA), calle 47 N° 257, La Plata.

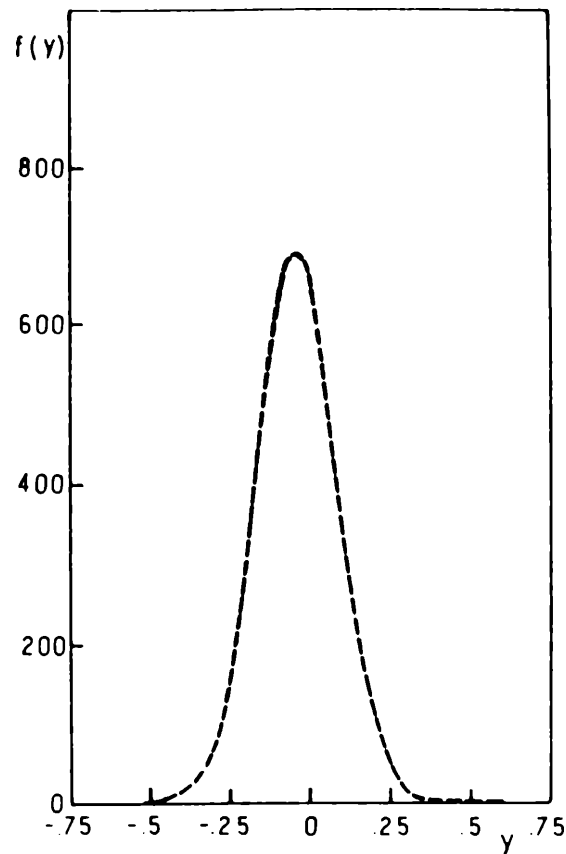


irradiada con un tubo de Cu de un espectrómetro Philips (Generador FW 1737/10 y goniómetro FW 1050/70).

Se midieron 40 puntos equiespaciados en  $\Delta 2\theta = 0.06$ , con los que se calcularon los momentos de orden cero, uno y dos. Por medio de estos momentos se obtuvieron los multiplicadores  $\lambda_1 = -0.9402$  y  $\lambda_2 = 683.66$  con los que se construyó la representación  $f(x)$  mostrada en la figura III.4. Esto nos permitió predecir el resto de los momentos dentro del error con que pueden calcularse a partir del perfil experimental.



**Fig. III.3:** a) Perfil experimental de la línea 220 para una hoja de Cu ( $x=\theta - \theta_0$  con  $\theta_0 = 37.13^\circ$ .) b) Perfil predicho con la representación hallada.



**Fig. III.4:** Distribución construida con dos multiplicadores  
( $y=\theta - \theta_0$  con  $\theta_0 = 37.13^\circ$ )

En la figura III.3.b se traza el perfil  $f_1^2$  predicho puntualmente por medio de  $f(x_k)$ . Como se ve en dicha figura, las predicciones de las colas del perfil se alejan de las conocidas más de lo que sería deseado. Según el método iterativo que propusimos, para mejorar estas predicciones deberíamos incorporar como información de partida un nuevo momento. En este caso, no es posible hacer esto debido a que se reproducen dentro del error

experimental todos los momentos mayores que  $10^{-18}$ . Por lo tanto, para poder mejorar las predicciones se debería disminuir el error experimental. Teniendo en cuenta la simplificación realizada en la determinación del ancho instrumental, los resultados deben considerarse dentro de esta aproximación.

### III.5 DISTRIBUCION DE TAMAÑOS CRISTALINOS EN ARCILLAS<sup>10</sup>

Despreciando los efectos instrumentales, el ensanchamiento del diagrama de difracción que producen las arcillas en estado puro, se debe a que las mismas se encuentran formadas por granos de muy pequeño espesor. Puesto que cada grano difracta independientemente, la intensidad registrada en un ángulo de difracción particular, contiene contribuciones de las intensidades producidas por granos de diferentes tamaños.

Consideremos una muestra en polvo, orientada según una dirección perpendicular al plano (001). Si  $N$  es el número de capas en esa dirección y  $d$  el espaciado interplanar, el espesor de cada grano será:  $e=Nd$ . La intensidad en el ángulo de difracción  $\theta$ , luego de ser corregida por factor de estructura y de Lorentz-polarización<sup>11</sup> es:

$$I(\theta) = A \langle \hat{G}_\theta \rangle$$

donde:

$$\hat{G}_\theta = \sum_N \langle N \rangle G(N, \gamma) \langle N \rangle$$

siendo:

$G(N, \gamma) = \frac{\text{sen}^2 N\gamma}{\text{sen}^2 \gamma}$ , la interferencia que produce un grano de  $N$  capas en el ángulo de difracción  $\theta$  relacionado con  $\gamma$  a través de la relación:

$$\gamma = 2\pi d \text{sen}\theta/\lambda$$

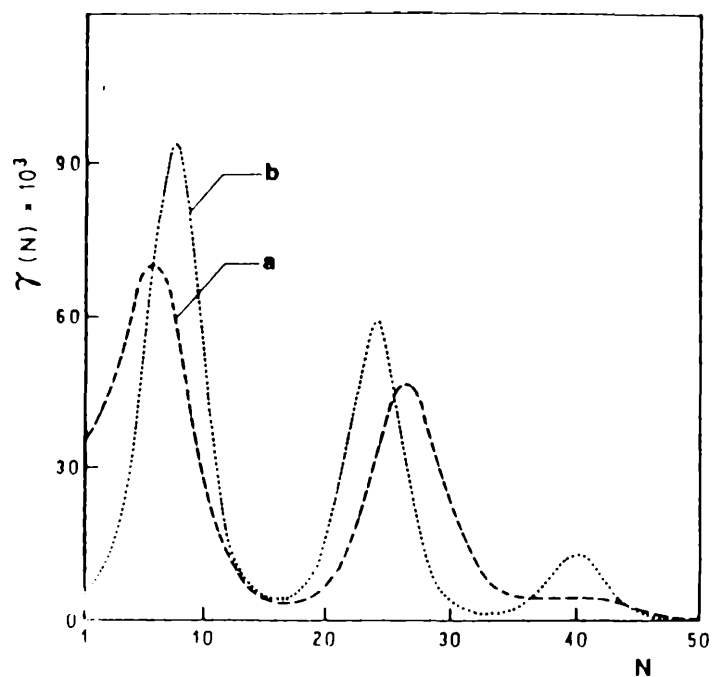
Llamando  $Y(N)$  a la distribución de tamaños:

$$I(Y) = A \sum_N Y(N) G(N, Y)$$

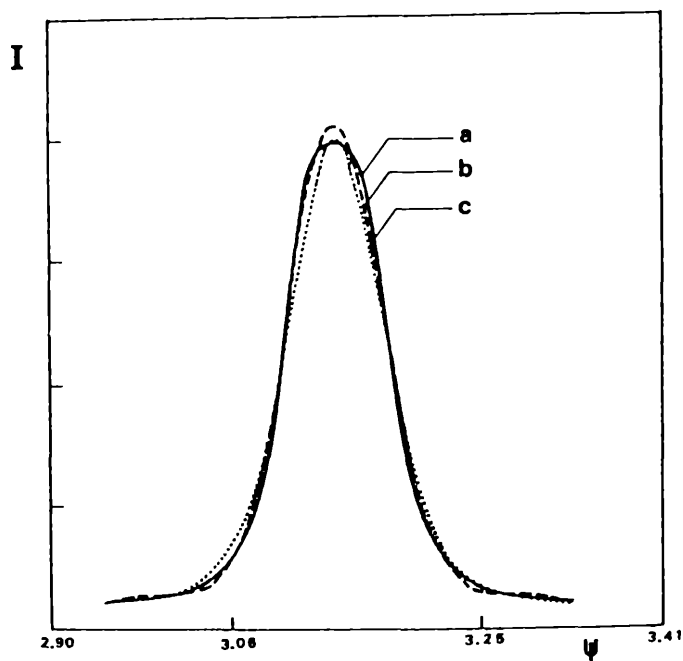
Por lo tanto, de la intensidad de difracción puede obtenerse la distribución  $Y(N)$ .

La curva de la figura III.5.a muestra la distribución de tamaños de una arcilla denominada caolinita<sup>3,12</sup>, obtenida con los cuatro primeros momentos del primer orden de difracción representado en la curva III.6.a. Los valores de intensidad predichos con esta distribución de tamaños, se ilustran en la curva III.6.b.

La distribución de tamaños obtenida con cuatro medidas puntuales de intensidad se presentan en la curva III.5.b mientras que la curva III.6.c muestra el pico de difracción que dicha distribución predice.



**Fig. III.5:** distribuciones de tamaños calculadas en la reflexión (001) de la caolinita: a) Utilizando los primeros cuatro momentos del pico de difracción. b) Utilizando cuatro datos independientes de intensidades.



**Fig. III.6:** Intensidades de difracción en la reflexión (001) de la caolinita. a) Pico experimental. b) Calculado con la distribución de tamaños de la curva III.5.a. c) Calculada con la distribución de tamaños de la curva III.5.b.

### III.6 CONCLUSIONES

Los ejemplos desarrollados en este capítulo muestran que incorporar el criterio de la TI como método de inversión, conduce a resultados sorprendentes. En la comparación con otros métodos se ve como, partiendo de mucho menos información, es posible obtener representaciones más satisfactorias.

En todos los casos presentados, hemos comparado las predicciones con abundante información disponible, alcanzando la convergencia deseada a partir de no más de cuatro medidas.

## REFERENCIAS

1. D.M.A. GUERIN, A.G. ALVAREZ, L.E. REBOLLO NEIRA, A. PLASTINO, R.D. BONETTO, Acta Cryst. A 42, 30-35(1985)
2. J. NUNEZ, L. REBOLLO NEIRA, A. PLASTINO, R. BONETTO, D. GUERIN, G. ALVAREZ, Deconvolution problems and information theory (enviado a publicar a Acta Cryst.)
3. F. RODRIGUEZ TRELLES, V Seminario de Análisis por Técnicas de Rayos X - La Falda - Córdoba (1985)
4. D.L. PHILLIPS, J. ACM, 9, 84 (1962)
5. S. TOWNEY. J. ACM, 10, 97 (1963)
6. G. BACKUS, F. GILBERT, Philos. Trans. Roy. Soc. London, A266, 123 (1970)
7. B.J. CONRATH, Inversion Methods in Atmospheric Remote Sounding. Deepak (1977)
8. M.T. CHAHINE, J. Atmos. Science, 27, 980, (1970)
9. M.T. CHAHINE, Inversion Methods in Atmospheric Remote Sounding, Deepak (1977)
10. R. BONETTO, D. GUERIN, G. ALVAREZ, L. REBOLLO NEIRA, A. PLASTINO, V Seminario de Análisis por Técnicas de Rayos X. La Falda - Córdoba (1985)
11. H.P. Klug, L.E. ALEXANDER, X-Ray Diffraction procedures, Jhon Wiley and Sons, New York (1974)



## **CAPITULO IV**

### **SENALES ESTADISTICAS**

Hasta el momento, hemos supuesto conocida la base donde una señal admite una representación positiva. Partiendo de esto, desarrollamos un formalismo y una metodología para encontrar dicha representación. En cierto sentido, hemos tratado problemas previamente planteados en la forma requerida para aplicar el método de inversión que proponemos.

En este capítulo particularizaremos un poco más el conjunto de señales consideradas, incorporándoles un determinado contenido. Más concretamente: analizaremos señales que contengan información sobre la distribución de estados de un sistema físico y las llamaremos "señales estadísticas". En este caso, sugerimos un modo general de proceder para encontrar la buena base de representación. Ejemplificaremos este procedimiento, aplicándolo a la determinación de secuencia de apilamiento de materiales interestratificados de dos componentes.

#### IV.1 DEFINICION DE SEÑALES ESTADISTICAS

Supongamos que interactuamos con un sistema físico, por medio de una señal  $|I\rangle$ , que llamaremos de entrada, la cual contiene y transmite información perfectamente conocida. Debido a la interacción, el sistema modificará esta señal transformándola en una señal  $|f\rangle$  que llamaremos respuesta. El proceso que acabamos de describir puede llevarse a un lenguaje matemático, del siguiente modo:

$$\hat{W} |I\rangle = |f\rangle \quad (IV.1)$$

Donde el operador  $\hat{W}$  representa la acción del sistema sobre la señal de entrada. Si conociéramos completamente el sistema y supiéramos de que modo actúa sobre la señal  $|I\rangle$ , entonces podríamos decir cual es la señal de salida sin necesidad de realizar el experimento. De lo contrario, podemos extraer información desconocida, estudiando la señal respuesta  $|f\rangle$ . Por ejemplo, por medio de un experimento adecuado, podemos decir que dos sistemas  $W_1$  y  $W_2$  se encuentran en estados diferentes si  $|f_1\rangle \neq |f_2\rangle$ .

El problema que aquí atacaremos es el de determinar la distribución de estados a partir de la señal  $|f\rangle$ . Este problema abarca, tanto el caso de un sistema que se encuentre en alguno de sus posibles estados de acuerdo a una ley probabilística, como el de un sistema compuesto por sub-sistemas de las mismas características pero en diferentes estados. No haremos distinción entre estos dos casos, puesto que ambos se caracterizan por lo que de ellos

nos falta conocer: la distribución de estados.

Consideremos entonces un sistema, del cual sabemos cuales son sus posibles estados y que además, adoptando algún modelo, podemos proponer de que forma actuaría sobre una señal de entrada, estando en un estado particular  $W_k$ . Es decir, proponemos la señal respuesta  $| \alpha \rangle$  construida por medio de la ecuación

$$\hat{W}_k | I \rangle = | \alpha \rangle \quad (IV.2)$$

Si designamos  $p_k$  a la probabilidad de que el sistema se encuentre en el estado  $W_k$ , el operador  $\hat{W}$  puede escribirse:

$$\hat{W} = \sum_k p_k \hat{W}_k \quad (IV.3)$$

Llamaremos "señal estadística" a la respuesta  $| f \rangle$  del sistema descrito, de manera que la misma pertenecerá a la imagen de la aplicación:

$$\hat{W} | I \rangle = \sum_k p_k | \alpha \rangle = | f \rangle \quad (IV.4)$$

Es decir, la representación de la señal  $| f \rangle$  respecto de la base  $| \alpha \rangle$  nos dará la distribución que estamos buscando. Para hacer esta asociación, suponemos que cada ket  $| \alpha \rangle$  respuesta de un estado determinado del sistema, es elemento de una base ortogonal del espacio de Hilbert.

Una medida sobre  $| f \rangle$  representable por un funcional lineal  $l$ , da cuenta de que sólo podemos conocer en valor medio el resultado de la misma medida sobre las señales  $| \alpha \rangle$ :

$$I_1(|f\rangle) = \sum_I p_i I_1(|\alpha\rangle)$$

$$f_i = \sum_I p_i L_{iI} \tag{IV.5}$$

Utilizando el formalismo desarrollado en el capítulo II, (IV.5) puede expresarse:

$$f_i = \text{TR}(\hat{f} \hat{L}_i) = \langle \hat{L}_i \rangle \tag{IV.6}$$

Por la propiedad de ortogonalidad, los elementos de la base  $\{|\alpha\rangle\}$  son autoestados del operador  $\hat{L}_i$ :

$$\hat{L}_i = \sum_I |\alpha\rangle L_{i\alpha} \langle \alpha| \tag{IV.7}$$

Cuando estos autoestados constituyan un conjunto completo, por analogía con la mecánica cuántica<sup>1</sup>, a los operadores  $\hat{L}_i$  les daremos el nombre de OBSERVABLES. Todo resultado de medir dicho observable es uno de sus autovalores y recíprocamente, todo autovalor es un posible resultado de esa medida, efectuada sobre una señal respuesta de un sistema en un estado particular. Pero únicamente podemos decir que un observable "tiene un valor particular" cuando estemos completamente seguros de que el resultado de medirlo es siempre ese valor particular. En caso contrario, sólo disponemos de su valor medio.

## IV.2 EL PROBLEMA DEL INTERESTRATIFICADO

El término interestratificado describe estructuras de filosilicatos, en las que dos o más capas ocurren en una secuencia vertical, esto es, a lo largo de una línea normal al plano (001). Las capas de filosilicatos están fuertemente unidas internamente, pero débilmente unidas unas a otras<sup>2</sup>. Luego, cada capa se aproxima a una molécula unidimensional en la dirección de apilamiento, y a un cristal bidimensional en las direcciones *a* y *b*. La superficie basal de diferentes clases de capas son geoméricamente muy similares y consisten de láminas de oxígenos o iones hidrógenos en un arreglo casi hexagonal.

La existencia de arcillas interestratificadas se conoce desde hace muchos años<sup>3,4</sup> y ha sido documentada por un estudio de Weaver (1956)<sup>5</sup>.

Aunque existen interestratificados de más de dos componentes, aquí nos limitaremos a considerar sistemas que contengan sólo dos clases de capas. Llamaremos genéricamente a una capa de un tipo particular como "capa 1" y la identificaremos con un 1 asociándole un espaciado interplanar  $d_1$ . Diferenciaremos otra capa de ésta, asociándole un 2 y un espaciado interplanar  $d_2$ . De esta manera, podemos representar las diferentes secuencias de apilamiento. Dos ejemplos de secuencia regular pueden ser:

$$1:2:1:2\dots \quad , \quad 1:2:2:1:2:2:\dots \quad (IV.8)$$

Un ejemplo de secuencia irregular es:

La figura de difracción producida por tales muestras interestratificadas varía según el grado de desorden de las mismas y la cantidad de material de uno u otro tipo. Por lo tanto, de un diagrama de difracción por rayos X, es posible extraer información acerca de la distribución de secuencia de apilamiento de las capas. Para esto, es necesario proponer la respuesta de cada secuencia o configuración particular, frente a interacciones con radiación X, cuya longitud de onda  $\lambda$  es del orden de los espaciados interplanares  $d_1$  y  $d_2$ .

#### IV.3 TRATAMIENTO SEMICLASICO DE LA DISPERSION DE RAYOS X

Para describir la dispersión elástica de la radiación X con la materia, adoptaremos un tratamiento semiclásico. No transcribiremos en detalle el desarrollo de este tratamiento, debido a que el mismo se encuentra en varios textos. En particular, los pasos que aquí omitimos, están detallados en ref. 6. Nos limitaremos entonces a hacer una breve reseña de las consideraciones que supone este tipo de tratamiento.

La descripción semiclásica considera al átomo aislado descrito cuánticamente, mientras que a la radiación X incidente se la trata clásicamente. La interacción es considerada como una perturbación que altera levemente los niveles estacionarios del átomo, y a la función de onda perturbada se la interpreta como describiendo un dipolo oscilante, que emite la radiación X dispersada.

Para encontrar el campo eléctrico asociado a la dispersión elástica, seguiremos el siguiente esquema:

- Por simplicidad tomaremos el caso de un átomo con un solo electrón, ya que esto no implica perder la esencia de la descripción.
- Necesitamos considerar sólo la interacción de la radiación con el electrón, porque al ser la masa del núcleo mucho mayor, su contribución a la dispersión de rayos X es despreciable.
- El electrón interactúa con el núcleo a través del potencial electrostático  $\phi$  y con el campo electromagnético a

través del potencial vectorial  $\mathbf{A}$ .

- Despreciando el término cuadrático en  $\mathbf{A}$ , el hamiltoniano toma la forma:

$$\hat{H} = \hat{H}_0 + \hat{H}' = (p^2/2m + e\phi) + (ie\hbar/mc) \vec{A} \cdot \vec{\nabla} \quad (\text{IV.10})$$

$\hat{H}'$  es la perturbación introducida por el campo incidente.

La función de onda puede ser calculada por medio de la teoría de perturbaciones<sup>7,8</sup>. El desarrollo perturbativo admite un truncamiento a primer orden como consecuencia de que la energía del fotón incidente,  $E = \hbar\omega$ , es mucho mayor que la diferencia de energía entre los niveles electrónicos y la amplitud del campo eléctrico incidente es pequeña. Por lo tanto, la función de onda que resulta es:

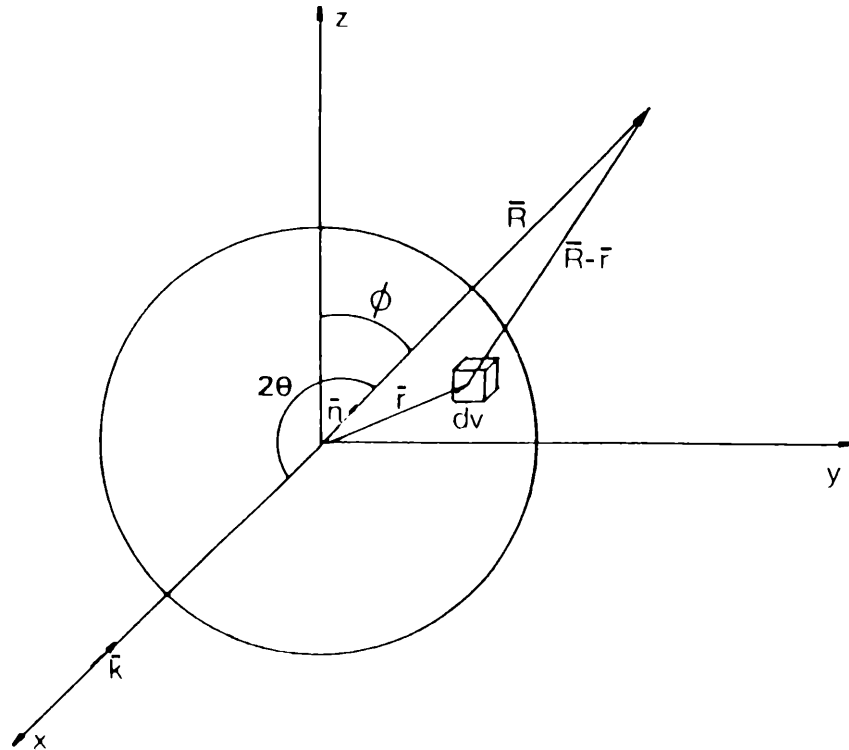
$$\Psi(r, t) = \Psi_n \exp(-i E_n t/\hbar) \quad (\text{IV.11})$$

siendo  $n$  el número que caracteriza el estado no perturbado. Con esta función de onda se calcula la densidad de corriente  $\vec{J}$ :

$$\vec{J} = i(e\hbar/2m) (\Psi \nabla \Psi^* - \Psi^* \nabla \Psi) - e^2/mc \vec{A} |\Psi|^2 \quad (\text{IV.12})$$

La radiación dispersada por el átomo, se obtiene considerando a esta densidad de corriente como fuente de ondas electromagnéticas secundarias, que obedecen las leyes del electromagnetismo clásico.





Considerando la geometría de la figura, la contribución de la corriente  $\vec{J}$  en  $dV$  al campo eléctrico de la onda dispersada, a la distancia  $R$  en que se encuentra el detector es:

$$d\vec{E} = 1/(c^2 R) \left\{ (d\vec{J}/dt \cdot dV \times \vec{n}) \times \vec{n} \right\} \Big|_{t'=t-(|\vec{R}-\vec{r}|)/c} \quad (\text{IV.13})$$

(IV.13) debe evaluarse al tiempo retardado  $t'=t-(|\vec{R}-\vec{r}|)/c$  donde  $|\vec{R}-\vec{r}|/c$  es el retraso de la contribución a la onda dispersada correspondiente al elemento de volumen  $dV$  en  $\vec{r}$ .

La expresión que se obtiene para  $d\vec{E}$  es la siguiente:

$$d\vec{E} = Q (\vec{E}_0 \times \vec{n}) \times \vec{n} \exp(i\omega R/c + i\omega t + i(\vec{k}' - \vec{k}) \cdot \vec{r}) |\gamma_n|^2 dV \quad (\text{IV.14})$$

siendo  $Q = e^2 / (R\omega c^2)$  y  $\vec{n}$  el versor en la dirección de dispersión, que es la misma que la de  $\vec{k}'$ .

El término  $(\vec{E}_0 \times \vec{n}) \times \vec{n}$ , depende de la polarización de la radiación incidente.

Llamando:

$$p(\theta) = Q |(\vec{E}_0 \times \vec{n}) \times \vec{n}| \quad (\text{IV.15})$$

reescribiremos (IV.14) en forma más compacta:

$$d\vec{E} = p(\theta) \exp(i\omega R/c + i\omega t + i(\vec{k}' - \vec{k}) \cdot \vec{r}) |\gamma_n|^2 dV \vec{e} \quad (\text{IV.16})$$

donde el versor  $\vec{e}$  indica la dirección  $(\vec{E}_0 \times \vec{n}) \times \vec{n}$ .

Esta expresión se interpreta\* como la contribución elemental al campo eléctrico de la onda dispersada, debida a una nube electrónica de densidad  $\rho(\vec{r}) = |\gamma_n|^2$ .

#### IV.4 DISPERSION POR UNA CAPA

Para poder resolver el problema de la dispersión de radiación X por arcillas interestratificadas, necesitamos conocer el campo eléctrico de la onda dispersada por cada capa de material uniforme. A partir de:

$$d\vec{E} = p(\theta) \exp(i\omega R/c + i\omega t + i(\vec{k}' - \vec{k}) \cdot \vec{r}) f(\vec{r}) dV \vec{e} \quad (\text{IV.17})$$

podemos calcular la contribución de una capa, efectuando la integral de (IV.17) sobre el volumen  $V_c$  de la capa.

Si  $d$  es la constante de red del material considerado y  $A$  el área de la capa, entonces:

$$V_c = Ad$$

y

$$\vec{E}_c = p(\theta) \exp(i\omega R/c + i\omega t) \int_{V_c} \exp(i(\vec{k}' - \vec{k}) \cdot \vec{r}) f(\vec{r}) dV \vec{e} \quad (\text{IV.18})$$

En un grano constituido por apilamiento de diferentes capas, podemos individualizar una de ellas fijando la coordenada  $z$ . Con esta coordenada fija, la densidad de carga es periódica en las otras dos y podemos desarrollarla en serie doble de Fourier:

$$f(x, y, z) = \sum_{h=-\infty}^{\infty} \sum_{k=-\infty}^{\infty} C_{hk}(z) \exp(-2\pi i (hx/a + ky/b)) \quad (\text{IV.19})$$

La ortogonalidad de la base nos permite obtener  $C_{hk}(z)$ :

$$C_{hk}(z) = 1/A_{cu} \int_{A_{cu}} f(\vec{r}) \exp(2\pi i (hx'/a + ky'/b)) dA \quad (\text{IV.20})$$

donde  $A_{cu}$  es el área construida sobre los vectores base  $\vec{a}$  y  $\vec{b}$  de la celda unitaria.

Reemplazando (IV.19) en (IV.18), resulta:

$$E_c = p(\theta) \exp(iwR/c+iwt) \sum_{hk} \int_{V_c} C_{hk}(z) \exp(-2\pi i (hx/a + ky/b + i(\vec{k}'-\vec{k}) \cdot \vec{r})) dV \vec{e} \quad (IV.21)$$

Evaluando la integral en x e y, ésta se anula a menos que se cumplan las condiciones de Laue:

$$2\pi h/a = k_x - k'_x \quad (IV.22)$$

$$2\pi k/b = k_y - k'_y$$

en cuyo caso la integral toma el valor A.

De todos los términos de la suma que aparece en (IV.21), sólo son distintos de cero aquellos que verifican las condiciones (IV.22). Luego:

$$E_c = p(\theta) \exp(iwR/c+iwt) A \int C_{hk}(z) \exp(i(\vec{k}'-\vec{k}) \cdot \vec{r}) dz \quad (IV.23)$$

Introduciendo en (IV.23) la expresión (IV.20) para  $C_{hk}$ , obtenemos:

$$E_c = p(\theta) \exp(iwR/c+iwt) A/ab \int_{V_c} f(r) \exp(2\pi i (hx/a + ky/b)) \exp(i(k'_z - k_z)z) dV \quad (IV.24)$$

$\Delta k_z = k'_z - k_z$ , depende del ángulo de difracción  $\theta$ .

Como  $|\vec{k}| = |\vec{k}'|$  (dispersión elástica), resulta:

$$\Delta k_z = 2 |\vec{k}| \sin\theta = 4\pi/\lambda \sin\theta \quad (IV.25)$$

y la integral en (IV.24) adopta la forma:

$$\int_{V_{cu}} f(\vec{r}) \exp(2\pi i (hx/a + ky/b + 2\text{sen}\theta z/\lambda)) dV \quad (\text{IV.26})$$

que se conoce en Cristalografía con el nombre de "factor de estructura". Estas son cantidades que conocemos en módulo y fase, para los materiales tratados. Como las experiencias se realizan sobre muestras orientadas según la dirección  $z$ , el valor de  $h$  y  $k$  en (IV.26) es cero. Además, el área basal de diferentes capas es muy similar una de otra<sup>2</sup> al igual que las constantes de red  $a$  y  $b$ . Esto nos permite considerar que el término  $A/ab$  en (IV.24) tiene aproximadamente el mismo valor  $C$ , cualquiera sea la capa que se esté tratando:

$$C = A/ab$$

Representaremos el factor de estructura dado por (IV.26) como un número complejo cuyo módulo  $F(j)$  y fase  $\phi$  dependen de  $\theta$ :

$$F(j)\exp(i\phi_j) = \int_{V_{cu}} f(r)\exp(4\pi i \text{sen}\theta z/\lambda) dV \quad j=1,2 \quad (\text{IV.27})$$

El subíndice  $j$  indica si se trata de una capa del tipo 1 o de una capa del tipo 2. De la misma manera, la contribución de cada capa al campo eléctrico será:

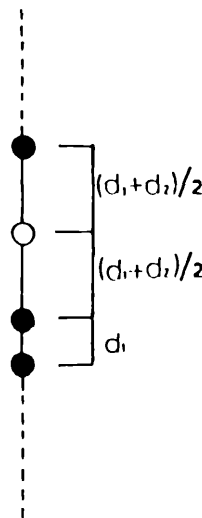
$$E_{c_j} = C p(\theta) \exp(i\omega R/c + i\omega t) F(j) \exp(i\phi_j) \quad (\text{IV.28})$$

Esta ecuación nos dice que es posible reemplazar una capa  $j$ , por un centro emisor de ondas secundarias con amplitud  $A_j(\theta) = C p(\theta) F(j)$  y una fase adicional  $\phi_j$  que proviene del factor de estructura.

#### IV.5 DIFRACCION DE N CAPAS INTERESTRATIFICADAS

Hemos visto que el campo asociado a la dispersión de una capa, puede interpretarse como producido por un centro dispersor puntual. Ubicaremos dichos puntos en posiciones que equidisten de las caras de las capas, representándolos con un punto lleno (●) o con un punto vacío (○), según se traten de centros de tipo 1 ó 2 respectivamente.

Así, la configuración 1:2:1:1 queda esquematizada en término de los centros dispersores del siguiente modo:

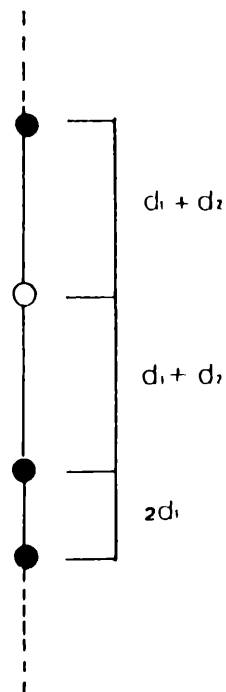


siendo  $d_1$  y  $d_2$  las constantes de red de las correspondientes capas

La contribución al campo eléctrico de cada centro dispersor está dada por la ecuación (IV.28). Debe notarse que existe un tiempo de retardo  $\Delta t$  entre dos centros dispersivos consecutivos:

$$\Delta t = (d_1+d_2) \sin\theta / (2c)$$

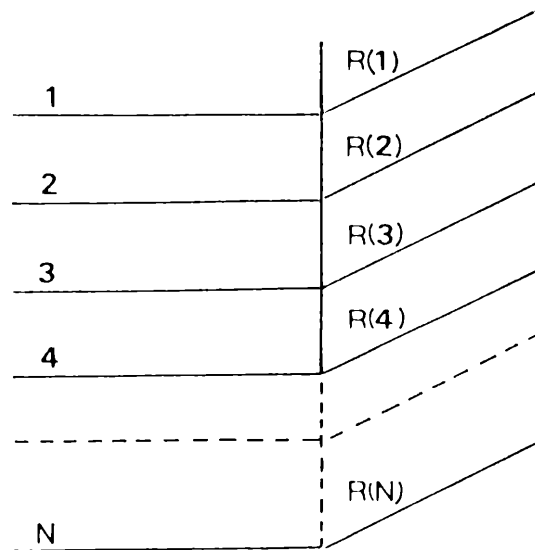
Como lo que importa es el retardo total  $\Delta t_r$  hasta llegar al detector y  $\Delta t_r = 2\Delta t$ , todavía es posible plantear una situación equivalente a la de la figura anterior, considerando a los centros como "centros emisores de ondas" separados una distancia  $(d_1+d_2)$ :



La señal respuesta que se producirá por dispersión de una dada configuración de  $N$  capas, dependerá del número de capas de cada tipo  $(N_1, N_2)$  y de la forma particular en que dichas capas se distribuyen. Es claro entonces que la base que representan estas señales respuesta, es la base del espacio de configuraciones, siendo cada secuencia de apilamiento una señal particular.

Para describir las distintas configuraciones usaremos un índice  $(i)$  que va de 1 a  $N$  y que rotula un nivel particular. Este nivel podrá estar ocupado por un centro (al que

indistintamente también llamaremos partícula) de tipo 1 o de tipo 2.



Según (IV.28) el campo eléctrico producido por una partícula que ocupa el nivel  $i$  es:

$$E(i) = C p(\theta) F(i) \exp(i\phi_i + i\omega t + i\omega R(i)/c) \quad (IV.29)$$

$R(i)$  es la distancia del centro  $i$  al detector.  $F(i)$  es el módulo del factor de estructura del centro que ocupa el nivel  $i$ . Este puede tomar los valores  $F(i) = F_1$  si el sitio está ocupado por una partícula 1, o  $F(i) = F_2$  si el sitio lo ocupa una partícula 2. Lo mismo ocurre con las fases  $\phi_i$  que pueden valer  $\phi_1$  o  $\phi_2$ .

Para obtener el campo eléctrico producido por  $N$  centros, hay que sumar las contribuciones de cada uno de ellos. Este tendrá la siguiente expresión general:

$$E_N = \sum_{i=1}^N E(i) = C p(\theta) \exp(i\omega t) \sum_{i=1}^N F(i) \exp(i\phi_i + i\omega R(i)/c) \quad (IV.30)$$



#### IV.6 SEÑALES RESPUESTA DE SISTEMAS INTERESTRATIFICADOS EN EL FORMALISMO NUMERO DE OCUPACIONES

Considerando fijo el número  $N$  de capas existen  $2^N$  configuraciones posibles. Sin embargo, en principio, el número de capas puede ser diferente para diferentes granos. Por esta razón, el espacio en que el problema estará bien representado, es un espacio de Hilbert extendido, en el cual el número de partículas es variable. Tal espacio se denomina espacio de Fock y para representarlo usaremos un formalismo similar al denominado número de ocupaciones. En este sentido, representaremos una señal respuesta de una configuración particular, como un estado base:

$$|i_1, i_2, \dots, i_N, \dots\rangle \quad (IV.31)$$

donde los índices  $i_1, i_2, \dots, i_N, \dots$  pueden tomar los valores 1 ó 2 según ese estado esté ocupado por una partícula de tipo 1 ó 2.

Por ejemplo, la configuración de la figura de la pag. 81 estará representada por un estado:

$$|1, 2, 1, 1\rangle \quad (IV.32)$$

Siguiendo con el formalismo, definiremos los operadores  $\hat{n}_i^1, \hat{n}_i^2$  tales que:

$$\hat{n}_i^1 |i_1, i_2, \dots, i_N, \dots\rangle = \begin{cases} |i_1, i_2, \dots, i_N, \dots\rangle & \text{si } i_i=1 \\ 0 & \text{si } i_i=2 \end{cases} \quad (IV.33)$$

$$\hat{n}_i^1 |i_1, i_2, \dots, i_1, \dots\rangle = \begin{cases} 0 & \text{si } i_1=1 \\ |i_1, i_2, \dots, i_1, \dots\rangle & \text{si } i_1=2 \end{cases} \quad (\text{IV.34})$$

Otra forma de escribir (IV.33) y (IV.34) es la siguiente:

$$\hat{n}_1^1 |i_1, i_2, \dots, i_1, \dots\rangle = (2-i_1) |i_1, i_2, \dots, i_1, \dots\rangle \quad (\text{IV.35})$$

$$\hat{n}_1^2 |i_1, i_2, \dots, i_1, \dots\rangle = (i_1-1) |i_1, i_2, \dots, i_1, \dots\rangle \quad (\text{IV.36})$$

En efecto, tomando el estado (IV.32), el nivel  $i_1$  da ocupado por una partícula 1, mientras que el  $i_2$  da ocupado por una partícula 2, como puede verse:

$$\hat{n}_1^1 |1, 2, 1, 1\rangle = (2-1) |1, 2, 1, 1\rangle = |1, 2, 1, 1\rangle$$

$$\hat{n}_1^2 |1, 2, 1, 1\rangle = (1-1) |1, 2, 1, 1\rangle = 0$$

$$\hat{n}_2^1 |1, 2, 1, 1\rangle = (2-2) |1, 2, 1, 1\rangle = 0$$

$$\hat{n}_2^2 |1, 2, 1, 1\rangle = (2-1) |1, 2, 1, 1\rangle = |1, 2, 1, 1\rangle$$

Los operadores  $\hat{N}_1$  (número de partículas 1) y  $\hat{N}_2$  (número de partículas 2) pueden definirse en término de los operadores  $\hat{n}_i^1$  y  $\hat{n}_i^2$ , de la siguiente manera:

$$\hat{N}_1 = \sum_{i=1}^M \hat{n}_i^1 \quad (\text{IV.37})$$

$$\hat{N}_2 = \sum_{i=1}^M \hat{n}_i^2 \quad (\text{IV.38})$$

El operador número total de partículas será obviamente:

$$\hat{N} = \hat{N}_1 + \hat{N}_2 = \sum_{i=1}^M \hat{n}_i^1 + \hat{n}_i^2 \quad (\text{IV.39})$$

#### IV.7 MEDIDAS COMO OBSERVABLES EN EL FORMALISMO NUMERO DE OCUPACIONES

Supondremos que una medida sobre una señal respuesta de una configuración particular, está dada por la proyección de dicha señal, sobre un determinado ángulo de difracción, digamos  $\theta$ . En este caso, tal medida corresponde a la intensidad de difracción en ese ángulo, producida por la configuración de que se trata.

Como se sabe, la intensidad de la onda electromagnética es proporcional al cuadrado de la amplitud del campo eléctrico  $EE^*$ . Llamando:

$$\tilde{G}_N(\theta) = E_N^* E_N \quad (IV.40)$$

y utilizando (IV.30), obtenemos:

$$\begin{aligned} \tilde{G}_N(\theta) = p^2(\theta) C^2 \{ & \sum_{i=1}^N F(i) \exp(i\theta_i + \\ & + i\omega R(i)/c) \} \left( \sum_{j=1}^N F(j) \exp(-i\theta_j - i\omega R(j)/c) \right) \end{aligned} \quad (IV.41)$$

$$\begin{aligned} = p^2(\theta) C^2 \{ & \sum_{i=1}^N F(i)^2 + \\ & + \sum_{i \neq j} F(i) F(j) \exp(i(\theta_i - \theta_j) + i\omega(R(i) - R(j))/c) \} \end{aligned} \quad (IV.42)$$

$$\begin{aligned} = p^2(\theta) C^2 \{ & \sum_{i=1}^N F^2(i) + \\ & + 2 \sum_{i < j} F(i) F(j) \cos(\theta_j - \theta_i + \omega(R(j) - R(i))/c) \} \end{aligned} \quad (IV.43)$$

donde  $p^2(\theta) = Q^2 I_0 ((1+\cos^2\theta)/2)$

Notando que  $w(R(j)-R(i))/c$  puede escribirse:

$$\begin{aligned} w(R(j)-R(i))/c &= (w/c) \sum_{s=i}^{j-1} (R(s+1)-R(s)) \\ &= (w/c) \sum_{s=i}^{j-1} \delta(s+1, s) \end{aligned} \quad (\text{IV.44})$$

donde  $\delta(s+1, s) = R(s+1) - R(s)$

e introduciendo:

$$\phi_j - \phi_i = \phi(j, i) \quad (\text{IV.45})$$

y

$$\tilde{G}_N(\theta) / (C^2 p^2(\theta)) = G_N(\theta) \quad (\text{IV.46})$$

en la ecuación (IV.43) resulta:

$$G_N(\theta) = \sum_{i=1}^N F^2(i) + 2 \sum_{i < j}^N F(i) F(j) \cos(\phi(j, i)) + \sum_{s=i}^{j-1} \delta(s+1, s) \quad (\text{IV.47})$$

Esta forma general de la intensidad producida por  $N$  partículas, adquirirá una forma particular, según sea la configuración que estas partículas adopten.

Recordemos que los  $F(i)$  pueden valer:

$$F(i) = \begin{cases} F1 \\ \\ F2 \end{cases} \quad (\text{IV.48})$$

los  $\phi_i$  adoptan los valores:

$$\delta_i = \begin{cases} \delta_1 \\ \delta_2 \end{cases} \quad (\text{IV.49})$$

de modo que:

$$\delta(j,i) = \begin{cases} 0 \\ \delta_1 - \delta_2 \\ \delta_2 - \delta_1 \end{cases} \quad (\text{IV.50})$$

mientras que los posibles valores de  $\delta(s+1,s)$  son:

$$\delta_{11} = 2d_1 w \text{ sen}\theta / c = 4\pi d_1 \text{ sen}\theta / \lambda$$

(si el nivel  $s+1$  y el  $s$  están ocupados por partículas del tipo 1)

$$\delta(s+1,s) = \delta_{22} = 4\pi d_2 \text{ sen}\theta / \lambda \quad (\text{IV.51})$$

(si los niveles están ocupados por partículas del tipo 2)

$$\delta_{12} = \delta_{21} = 2\pi(d_1 + d_2) \text{ sen}\theta / \lambda$$

(si cada nivel está ocupado por una partícula de diferente tipo)

Estas variables aparecerán tomando alguno de sus valores posibles, dependiendo de la configuración que se considere. Resulta entonces de utilidad definir los siguientes operadores:

$$\hat{F}_k(i) = F1_k \hat{n}_i^1 + F2_k \hat{n}_i^2$$

$$\hat{\theta}_k(j, i) = \theta1_k \hat{n}_j^1 + \theta2_k \hat{n}_i^2 - (\theta1_k \hat{n}_i^1 + \theta2_k \hat{n}_j^2) \quad (IV.52)$$

$$\begin{aligned} \hat{\delta}_k(s+1, s) = & \delta11_k \hat{n}_{s+1}^1 \hat{n}_s^1 + \delta22_k \hat{n}_{s+1}^2 \hat{n}_s^2 + \\ & + \delta21_k \hat{n}_{s+1}^2 \hat{n}_s^1 + \delta12_k \hat{n}_{s+1}^1 \hat{n}_s^2 \end{aligned}$$

El subíndice  $k$  indica que se ha fijado el ángulo de difracción en  $\theta_k$ . En término de los operadores definidos, una medida en este ángulo particular puede expresarse en forma proporcional a un operador  $\hat{G}_k$  dado por la expresión:

$$\hat{G}_k = \sum_{i=1}^N \hat{F}_k^2(i) + 2 \sum_{i < j}^N \hat{F}_k(i) \hat{F}_k(j) \cos(\hat{\theta}_k(j, i)) + \sum_{s=1}^{j-1} \hat{\delta}_k(s+1, s) \quad (IV.53)$$

Puesto que  $\hat{G}_k$  es una función analítica de operadores lineales en  $\hat{n}_i^1$  y  $\hat{n}_i^2$ , los estados base son autoestados de este operador, con autovalores dados por la ecuación (IV.47) evaluada en el ángulo  $\theta_k$  para cada configuración.

Veamos como ejemplo, como actúa  $\hat{G}_k$  sobre el estado  $|1, 2, 2\rangle$ :

$$\begin{aligned}
\hat{G}_k |1,2,2\rangle = & \left\{ \sum_{i=1}^3 (F1 \hat{n}_i^1 + F2 \hat{n}_i^2) \right\}^2 + \\
& + 2 \sum_{i < j}^3 (F1 \hat{n}_i^1 + F2 \hat{n}_i^2) (F1 \hat{n}_j^1 + \\
& + F2 \hat{n}_j^2) \cos(\theta 1_k (\hat{n}_j^1 - \hat{n}_i^1) + \\
& + \theta 2_k (\hat{n}_j^2 - \hat{n}_i^2)) + \\
& + \sum_{s=1}^{j-1} \delta 11_k \hat{n}_{s+1}^1 \hat{n}_s^1 + \delta 22_k \hat{n}_{s+1}^2 \hat{n}_s^2 + \\
& + \delta 21_k \hat{n}_{s+1}^2 \hat{n}_s^1 + \delta 21_k \hat{n}_{s+1}^1 \hat{n}_s^2 \} |1,2,2\rangle
\end{aligned}$$

$$\begin{aligned}
\hat{G}_k |1,2,2\rangle = & (F1_k^2 + F2_k^2 + F2_k^2 + 2F1_k F2_k \cos(\theta 2_k - \\
& - \theta 1_k + \delta 21_k) + 2F1_k F2_k \cos(\theta 2_k - \\
& - \theta 1_k + \delta 21_k + \delta 22_k) + \\
& + 2F2_k F2_k \cos(\theta 2_k - \theta 2_k + \delta 22_k) \} |1,2,2\rangle
\end{aligned}$$

Si designamos con la letra  $\nu$  al conjunto de índices necesarios para caracterizar una configuración de  $N$  partículas, a la identidad podremos expresarla como:

$$I = \sum_{k=1}^N \sum_{\nu=1}^{2^N} |\nu\rangle \langle \nu| \tag{IV.54}$$



Una medida  $f_n$  sobre la señal respuesta del sistema, en el ángulo de difracción  $\theta_n$  es proporcional al valor medio del operador  $\hat{G}_n$  dado por (IV.53):

$$f_n \sim \langle G_n \rangle \quad (IV.55)$$

#### IV.8 CONCLUSIONES

Para analizar señales estadísticas, hemos propuesto que una señal respuesta del sistema en un determinado estado es autoestado de un observable que representa una medida sobre la señal. Por lo tanto, manteniendo fija la señal de entrada  $|I\rangle$ , el conjunto de medidas lineales sobre la señal respuesta  $|f\rangle$  dan los valores medios de un conjunto de observables mutuamente conmutantes. Sin embargo, una medida sobre el sistema supone una determinada interacción y cabe preguntarnos si es posible extraer información por medio de medidas sobre señales que son respuesta de diferentes interacciones. Del formalismo que presentamos surge como respuesta que diferentes medidas son compatibles, si admiten una representación diagonal en la misma base. Si la interacción es fija, esto siempre ocurre. Si cambiamos la interacción, sólo existirá compatibilidad cuando las señales respuesta de diferentes interacciones formen una base común.

## REFERENCIAS

1. P.A.M. DIRAC, The Principles of Quantum Mechanics, Oxford University Press (1985)
2. G.W. BRINDLEY, G. BROWN, Cristal Structures of Clay Minerals and their X-Ray Identification, Mineralogical Society, (1980)
3. J.W. GRUNER, Am. Miner. 19, 557-578 (1934)
4. L.T. ALEXANDER, S.B. HENDRICKS, R.A. NELSON, Soil. Sci, 44, 237-239 (1939)
5. C.E. WEAVER, Am. Min. 41, 202-221 (1956)
6. R.W. JAMES, The Optical Principles of Diffraction of X-Ray, G. Bell and Sons. Ed, London (1954)
7. E. MERZBACHER, Quantum Mechanics, John Wiley and Sons. Inc. (1970)
8. A. MESSIAH, Mecánica Cuántica, Ed. Tecnos S.A. (1975)
9. L.D. LANDAU, E.M. LIFSHITZ, Teoría Clásica de Campos, Reverté (1966)

## **CAPITULO V**

### **ANALISIS DEL DIAGRAMA DE DIFRACCION DE ARCILLAS INTERESTRATIFICADAS**

En este capítulo se muestran los resultados que surgen de analizar los diagramas de difracción de rayos X de dos muestras interestratificadas de Ilita - Montmorillonita.

El operador estadístico se construye a partir de medidas puntuales y por medio de él se calcula la proporción en que se encuentran los dos materiales de la muestra.

## V.1 CARACTERISTICAS GENERALES DE LOS DOS COMPONENTES DE UNA MUESTRA DE ILITA - MONTMORILLONITA.

La montmorillonita es una arcilla que pertenece al grupo de las esmectitas. Estas arcillas tienen estructura en capas en las que oxianiones bidimensionales están separados por capas de cationes hidratados. Los átomos de oxígeno definen capas de sitios tetraedrales por encima y por debajo de una capa central de sitios octaedrales<sup>1,2</sup>.

Los miembros del grupo de las esmectitas son distinguibles por el tipo y localización de cationes en el armazón de oxígenos. La diferencia en la carga de las capas produce diferencias en las propiedades físicas y químicas. En la montmorillonita, la carga de las capas se origina por la sustitución de  $Al^{3+}$  octaedrales por  $Mg^{2+}$ . Esta sustitución origina un defecto de carga positiva, que se manifiesta como un exceso de carga negativa sobre la superficie de las láminas de arcilla. Esta carga negativa de la red, es compensada por cationes que son adsorbidos en los espacios interlaminares. En presencia de agua o vapor de agua, la montmorillonita se expande debido a la penetración de las moléculas de agua entre la separación de las celdas unitarias.

La montmorillonita que interviene como componente de las muestras interestratificadas que analizaremos se denomina "sódica", por ser cationes de Na los que se hallan como compensantes del defecto de carga en la red.

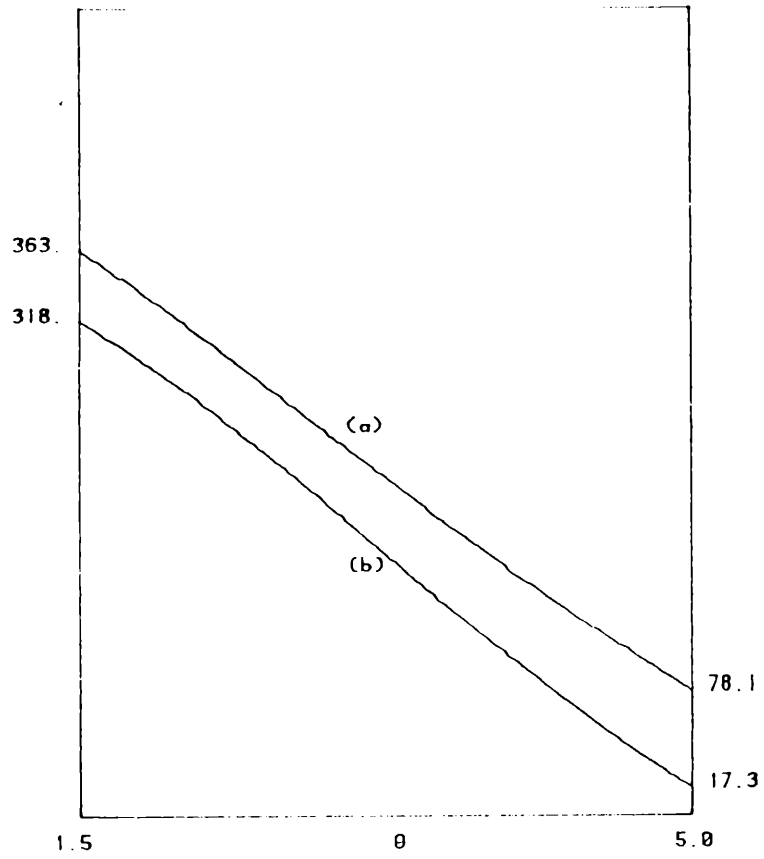
Las ilitas no poseen la capacidad de intercambiar cationes

ni de expandirse. En la illita que consideraremos, el exceso de carga de la superficie es neutralizada por cationes de K. Debido a que algunas illitas provienen de la montmorillonita, es muy difícil que éstas se encuentren en estado totalmente puro. En mayor o menor grado siempre están en algún estado de interestratificación.

En el capítulo IV hemos distinguido cada componente de una muestra interestratificada con los nombres genéricos de "tipo 1" y "tipo 2". Aquí le asignaremos el 1 a la montmorillonita y el 2 a la illita. Los correspondientes espaciados valen  $d_1 = 15.8 \text{ \AA}$  y  $d_2 = 10 \text{ \AA}$ .

Como consecuencia de poseer un centro de simetría, las fases del factor de estructura de una y otra arcilla sólo pueden valer 0 ó  $\pi$ . En el rango de ángulos en que aparecen los diagramas estudiados, ambas fases valen cero.

La figura V.1 muestra los módulos de los factores de estructura en función del ángulo de difracción. La curva a) representa el de la montmorillonita y fue obtenida por medio de los valores publicados en ref 3. La curva b) que corresponde a la illita, fue calculada a partir de las posiciones atómicas conocidas y que aparecen en ref 4.



**Fig. V.1:** Módulos de los factores de estructura en función del ángulo de difracción. Curva a)  $F(1)$  correspondiente a la montmorillonita. Curva b)  $F(2)$  correspondiente a la illita.

## V.2 DETERMINACION DEL GRADO DE INTERESTRATIFICACION DE DOS MUESTRAS DIFERENTES DE ILITA-MONTMORILLONITA.

En las figuras V.3.a y V.3.b pueden verse las medidas puntuales obtenidas por difracción de Rayos X de dos muestras policristalinas de ilita-montmorillonita\*.

Los datos graficados han sido corregidos por el factor de polarización  $(1+\cos^2 2\theta)/2$  y el factor de Lorentz  $L(\theta)$ . Este último factor es característico de la forma de recolectar los datos<sup>6</sup>. En este caso, por tratarse de muestras orientadas, toma la forma:

$$L(\theta) = 1/(\sin \theta \cos \theta)$$

Al producto de estos dos factores se lo conoce con el nombre de Factor de Lorentz-polarización,  $L_p(\theta)$ . El mismo se encuentra graficado en la figura V.2, dentro del rango en que se lo ha utilizado.

\* Las muestras fueron cedidas por el Dr. Mario Iñiguez del Centro de Investigaciones Geológicas (CIG) y fueron medidas por el Grupo de Física del Estado Sólido del CINDECA.



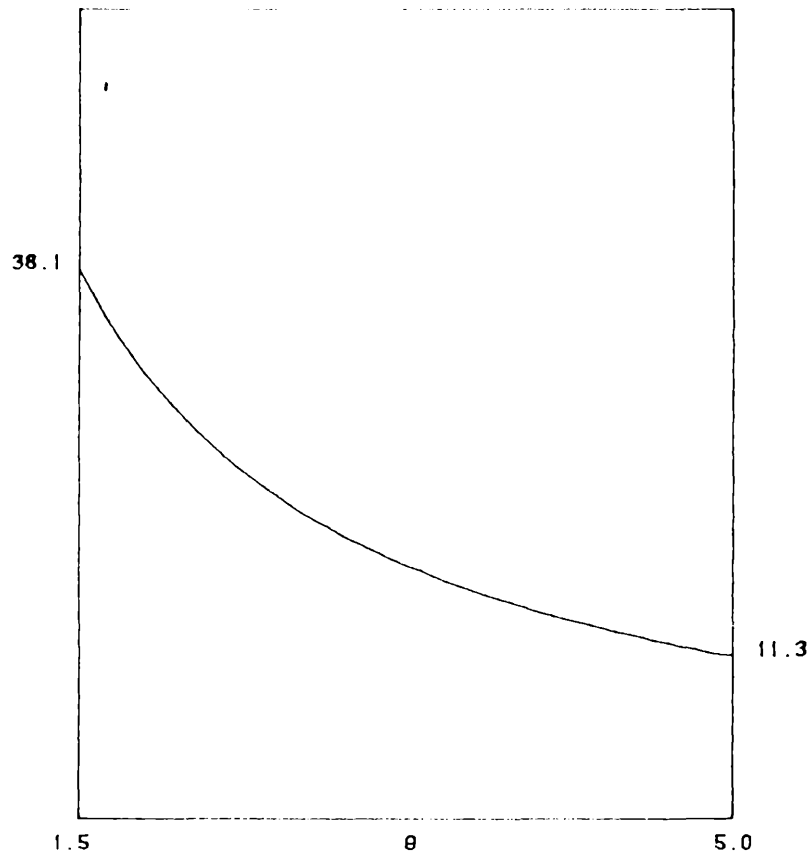


Fig. V.2 Factor de Lorentz-polarización  
 $L_p(\theta) = (1 + \cos^2 2\theta) / \sin 2\theta$

El operador estadístico que se obtiene a partir de las medidas puntuales  $\langle \theta_i | f \rangle$ , correspondientes a la intensidad de difracción está dado por:

$$\hat{P} = \exp\left(-\sum_{i=3}^N \lambda_i \hat{O}_i\right) / \text{TR}\left(\exp\left(-\sum_{i=3}^N \lambda_i \hat{O}_i\right)\right) \quad (\text{V.1})$$

donde:

$$\begin{aligned} \hat{O}_i = & \{ \langle \theta_i | f \rangle - \langle \theta_1 | f \rangle \} \hat{G}_2 - \{ \langle \theta_i | f \rangle - \langle \theta_2 | f \rangle \} \hat{G}_1 - \\ & - \{ \langle \theta_2 | f \rangle - \langle \theta_1 | f \rangle \} \hat{G}_i \end{aligned} \quad (\text{V.2})$$

Esta forma particular que adoptan los operadores  $\hat{O}_i$  es consecuencia de haber tomado una medida adicional para ajustar un fondo constante  $B$ . Por esta misma razón el factor de proporcionalidad  $A$  debe calcularse por medio de la siguiente expresión:

$$A = \{\langle \theta_2 | f \rangle - \langle \theta_1 | f \rangle\} / \{\langle \hat{G}_2 \rangle - \langle \hat{G}_1 \rangle\} \quad (V.3)$$

ella permite obtener  $B$  a través de:

$$B = \langle \theta_2 | f \rangle - A \langle \hat{G}_2 \rangle \quad (V.4)$$

Los operadores  $\hat{G}_i$  que intervienen en (V.2), (V.3) y (V.4) son los definidos por la ecuación (IV.53).

En la sección IV.7 del capítulo anterior, hemos designado con la letra  $\nu$  al conjunto de índices necesarios para caracterizar una configuración particular. Puesto que el número de configuraciones crece muy rápidamente con el número de partículas, para efectuar las trazas fue necesario considerar un número de partículas máximo que se fijó en 10. Se permitió superar este número sólo a las configuraciones en las que intervienen partículas del mismo tipo. La elección del valor 10 para el número máximo de partículas en estado interestratificado, se realizó teniendo en cuenta que la distribución de tamaño de granos de una montmorillonita en estado casi puro es despreciable, si el número de capas es superior a 10 (ref. 6,7). Efectuando este truncamiento, el número de configuraciones se redujo a 2100.

Para obtener convergencia a los 63 datos de la figura V.3.a dentro del error experimental, representado por las barras, fue necesario partir de 10 medidas puntuales para determinar 8 multiplicadores de Lagrange. Los 10 valores de las intensidades de partida son los siguientes:

$\theta_i$	$\langle \theta_i   f \rangle$
1.8	279
2.2	265
2.5	313
2.8	383
3.1	445
3.4	389
3.7	388
4.0	346
4.2	320
4.4	322

Por medio de estas intensidades es posible inferir las 53 restantes a través de la expresión:

$$f_i^* = A \langle \hat{G}_i \rangle + B \tag{V.5}$$

Como lo muestra la curva llena de la figura V.3.a, estas predicciones convergen a las medidas conocidas dentro del error con que éstas se conocen (3%).

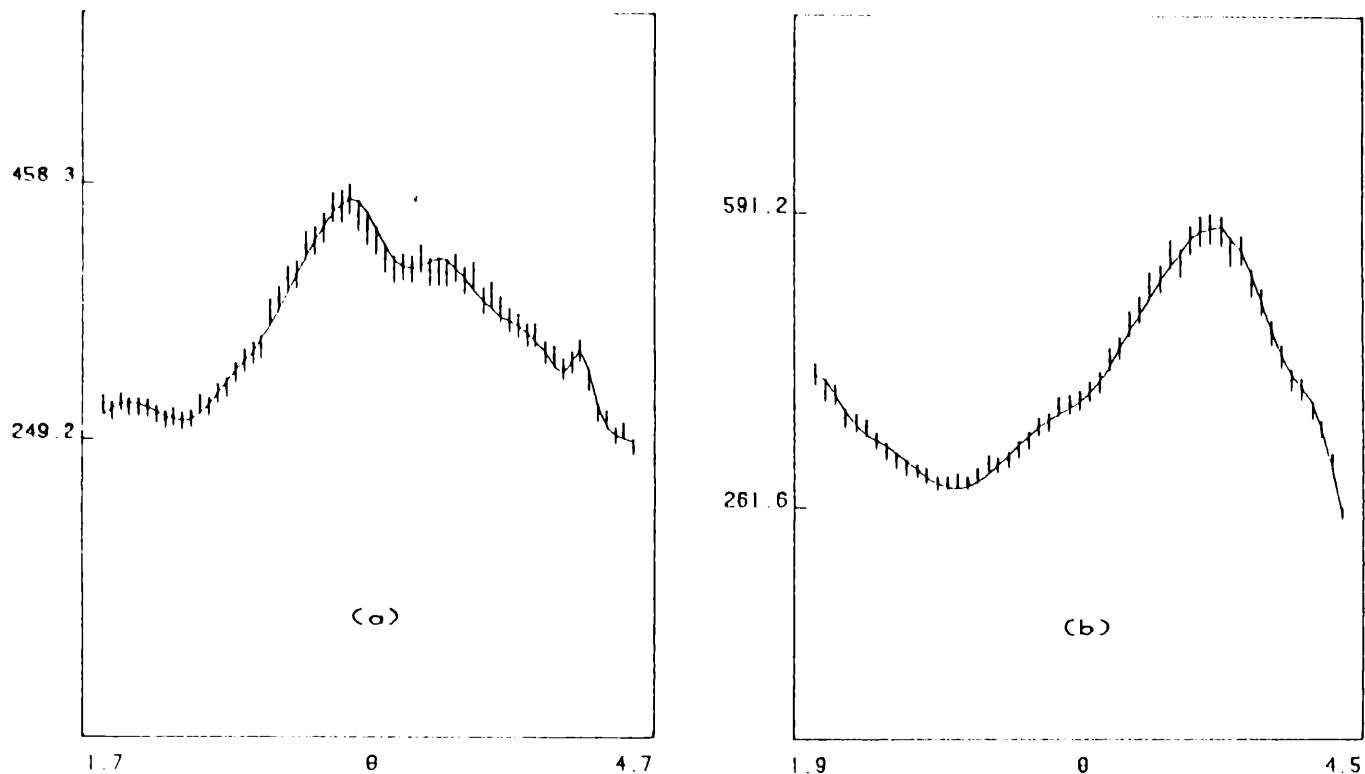


Fig. V.3: Intensidades corregidas por factor de Lorentz-polarización, las barras corresponden a un error del 3%.

- a) La curva continua representa las intensidades que se infieren estadísticamente con  $\hat{P}^0$ .
- b) La curva continua representa las intensidades que se infieren estadísticamente con  $\hat{P}^{11}$ .

Puesto que el operador estadístico es capaz de efectuar predicciones correctas, podemos utilizarlo para calcular la cantidad de material de cada tipo que está presente en la muestra:

$$\langle \hat{N}_1 \rangle = \text{TR}(\hat{P}^0 \hat{N}_1) \tag{V.6}$$

$$\langle \hat{N}_2 \rangle = \text{TR}(\hat{P}^0 \hat{N}_2) \tag{V.7}$$

Los valores porcentuales que se obtuvieron son los siguientes:

MONTMORILLONITA	30%
ILITA	70%

En la segunda muestra analizada se logró converger a los 53 datos de la figura V.3.b, por medio de 13 medidas, que se usaron para calcular 11 multiplicadores.

$\theta_i$	$\langle \theta_i   f \rangle$
1.9	412
2.2	337
2.5	288
2.8	310
3.0	354
3.2	383
3.4	441
3.6	517
3.8	571
4.0	549
4.2	433
4.4	352
4.5	254

Las 40 medidas restantes se calcularon por medio del operador  $\hat{P}^{11}$ . La curva de la figura V.3.b muestra la convergencia de las predicciones a los valores experimentales. En este segundo caso los valores porcentuales obtenidos son:

MONTMORILLONITA      20%  
 ILITA                      80%

En este último caso la distribución de configuraciones,  $\langle \hat{\rho} | \hat{\rho} \rangle$ , indica que existen unos pocos estados más densamente poblados que el resto. A continuación mostramos las componentes más significativas.

CONFIGURACION	$f_i$
1	0.146
2	0.108
122	0.078
221	0.078
222	0.051
22122	0.029
222222	0.027
12222	0.025
22221	0.025
22222222	0.021
2222	0.019
22222	0.015
11	0.014
12221	0.012
22	0.010
12222222	0.009
22222221	0.009
1222122	0.009
2212221	0.009

### V.3 ESTIMACION DE ERRORES

Para estimar la influencia de los errores con que se conocen las intensidades, en el cálculo del grado de interestratificación realizado en la sección anterior, llamaremos:

$I_i^o$  a las intensidades medidas en el ángulo  $\theta$ .

$I_i^y$  a las intensidades esperadas

$I_i^p$  a las intensidades predichas

Puesto que las medidas están realizadas dentro de un error del 3%, llamando  $\epsilon$  al factor 3/100, las intensidades esperadas pertenecen al siguiente intervalo:

$$I_i^o (1-\epsilon) \leq I_i^y \leq I_i^o (1+\epsilon) \quad (V.8)$$

Si  $K$  es el número de medidas conocidas, y  $M$  el número de medidas utilizadas para obtener convergencia, las intensidades predichas por medio del operador  $\hat{F}^n$  satisfacen:

$$I_i^o (1-\epsilon) \leq I_i^p \leq I_i^o (1+\epsilon) \quad i=1, \dots, k \quad (V.9)$$

Por lo tanto:

$$I_i^p / (1+\epsilon) \leq I_i^y \leq I_i^p / (1-\epsilon) \quad (V.10)$$

y:

$$I_i^p (1-\epsilon) / (1+\epsilon) \leq I_i^y \leq I_i^p (1+\epsilon) / (1-\epsilon) \quad (V.11)$$

Llamando  $\hat{F}_v$  al operador estadístico capaz de reproducir los valores esperados de la intensidad y  $A_v$  a la correspondiente constante de proporcionalidad:

$$A(1-\epsilon)/(1+\epsilon) \hat{F}^n \leq A_v \hat{F}^v \leq A(1+\epsilon)/(1-\epsilon) \hat{F}^n \quad (V.12)$$

Por ser tanto  $\hat{F}^v$  como  $\hat{F}^n$  operadores normalizados:

$$A(1-\epsilon)/(1+\epsilon) \leq A_v \leq A(1+\epsilon)/(1-\epsilon) \quad (V.13)$$

Luego:

$$(1-\epsilon)^2/(1+\epsilon)^2 \hat{F}^n \leq \hat{F}^v \leq (1+\epsilon)^2/(1-\epsilon)^2 \hat{F}^n \quad (V.14)$$

Multiplicando por los correspondientes operadores  $\hat{N}_1$  y  $\hat{N}_2$  y efectuando las trazas:

$$(1-\epsilon)^2/(1+\epsilon)^2 \langle \hat{N}_1 \rangle \leq \langle \hat{N}_1^v \rangle \leq (1+\epsilon)^2/(1-\epsilon)^2 \langle \hat{N}_1 \rangle \quad (V.15)$$

$$(1-\epsilon)^2/(1+\epsilon)^2 \langle \hat{N}_2 \rangle \leq \langle \hat{N}_2^v \rangle \leq (1+\epsilon)^2/(1-\epsilon)^2 \langle \hat{N}_2 \rangle \quad (V.16)$$

Estas desigualdades implican que los porcentajes del grado de interestratificación que hemos calculado, están afectados por un error del 12%.



#### V.4 CONCLUSIONES

A pesar de que los porcentajes de un mismo material que hemos encontrado en cada muestra sólo difieren en un 10%, los diagramas de difracción que éstas producen son notoriamente diferentes. Tal diferencia es atribuible a la desigualdad en las entropías de uno y otro caso. Los elementos de matriz del operador estadístico que nos permitió predecir la curva V.3.a, indican un elevado número de estados con similar probabilidad (por esta razón no presentamos tabla de valores), mientras que en la segunda muestra existen estados más densamente poblados que los demás.

Otro hecho significativo es que, aunque los máximos valores de intensidad ocurren en diferentes ángulos, ambas muestras contienen un elevado porcentaje de illita. Esto se debe a que el factor de estructura de la montmorillonita, en el ángulo donde ésta presenta su máximo de difracción en estado puro (aproximadamente  $2.85^\circ$ ), es cuatro veces mayor que el de la illita en el ángulo que la caracteriza (aproximadamente  $4.4^\circ$ ).

Así, del 70% de illita presente en una de las muestras, un 22% existente en estado puro es el que produce el pequeño pico que puede observarse en  $4.4^\circ$  (fig. V.3.a), mientras que el porcentaje de montmorillonita pura es despreciable, a pesar de que el máximo principal de la misma figura ocurre en  $3.1^\circ$ .

Con el fin de comprobar la validez del truncamiento en el

número de configuraciones, hemos realizado algunas pruebas, permitiendo hasta 12 capas interestratificadas, lo que corresponde a un número de configuraciones del orden de 8000. Los porcentajes encontrados no variaron más que en un 1%.

## REFERENCIAS

1. G.W. BRINDLEY, G. BROWN, *Cristal Structures of Clay Minerals and their X-Ray Identification*, Mineralogical Society (1980)
2. T. PINNAVIA, *Science*, Vol. 220, N° 4595, 365-371 (1983)
3. W.F. COLE, C.J. LANCUCKI, *Acta Cryst.*, 21, 836-838 (1966)
4. C.E. WEAVER and L.D. POLLARD, *The Chemistry of Clay Minerals*, Elsevier Sci. Publ. (1973)
5. H.P. KLUG, L. ALEXANDER, *X-Ray Diffraction Procedures*, Willey and Sons. (1974)
6. D.M.A. GUERIN, A.G. ALVAREZ, L.E. REBOLLO NEIRA, A. PLASTINO, R.D. BONETTO, *Acta Cryst.*, A42, 30-35 (1985)
7. R. BONETTO, D. GUERIN, G. ALVAREZ, L. REBOLLO NEIRA, A. PLASTINO, V Seminario de Análisis por técnicas de Rayos X, La Falda - Córdoba (1985)

## CONCLUSIONES GENERALES

Cuando decidimos adoptar el criterio que provee la TI para construir el operador estadístico, a partir de medidas sobre señales, lo hicimos basándonos en el prestigio del segundo postulado de la TI. Dicho prestigio lo atribuimos al hecho de que, cuando el postulado se utiliza en el tratamiento de otros problemas, los resultados que produce son resultados esperados. Ejemplo de ello es que permite reformular la Mecánica Estadística y la Teoría de Scattering, interpretar desde otro punto de vista resultados de la Mecánica Cuántica, etc.

Sin embargo, al obtener los primeros resultados que el método de inversión produce, comprobamos que ya no era necesario basarnos en el prestigio del postulado para presentar como óptima la solución que proponíamos. No necesitábamos fundarnos en confrontaciones anteriores porque eran nuestras propias medidas las que nos permitían confiar en la solución que provee el postulado.

En todos los casos tratados hemos logrado convergencia a la información disponible a partir de un número reducido de datos experimentales, mientras que el resto de los mismos, sirvió para comparar la validez de nuestras conjeturas.

El problema de determinar el grado de interestratificación en arcillas, nos lo planteamos como un verdadero desafío al método propuesto, por la gran dimensión que el mismo involucra. Sin embargo, aún en este caso hemos

podido verificar la bondad de la solución propuesta. Los resultados que presentamos en este trabajo contribuyen a aumentar el prestigio del que ya goza el postulado de la TI, y creemos que aún en los casos donde la información disponible sea escasa, tendremos derecho a decir que la solución propuesta es la mejor que podemos dar en base a la información que poseemos.

Rescatamos también el carácter predictivo de la representación que se obtiene adoptando la TI en el siguiente sentido:

Si la base de representación de una señal es conocida y el problema no involucra una dimensión demasiado elevada, son muy pocas las medidas por medio de las cuales es posible obtener una representación satisfactoria. Pero si la base no se conoce y se propone alguna, tentativa, la TI permite desechar bases incorrectas por dos motivos:

- Si la señal no admite representación positiva en la base elegida, las medidas contendrán información contradictoria y no podrán determinarse los multiplicadores de Lagrange correspondientes.

- Si admite tal representación pero no se logra predecir correctamente, salvo utilizando casi todas las medidas disponibles, la base puede considerarse inadecuada para representar el sistema estudiado.

## AGRADECIMIENTOS

Del Dr. Angel Plastino, he recibido la mejor de las direcciones a que pudiera aspirar. Además de sus valiosas enseñanzas, le agradezco muy especialmente que haya motivado, guiado y enriquecido mis iniciativas, a pesar de la modestia de las mismas.

Han participado en este trabajo: Carlos Esebbag, María Teresa Martín y Josué Núñez; trabajar junto a ellos ha sido particularmente grato. Agradezco además las contribuciones del resto de los integrantes del Grupo de Teoría Nuclear y la permanente colaboración de Graciela Jaen.

También son parte del trabajo el grupo de Física del Estado Sólido del CINDECA: Dr. Guillermo Alvarez, Rita Bonetto y Diego Guérin. Ellos han puesto la mejor de las voluntades para lograr que esta tarea interdisciplinaria se realice, en un ambiente en el que siempre me sentí muy cómoda. Mi agradecimiento se extiende a todo el personal del CINDECA, por la hospitalidad con que me recibieron.

En materia de computación recibí asesoramiento y apoyo de Silvia Gordillo, Catalina Mostaccio y Jorge Pouchou del Laboratorio de Computación y Cálculo Numérico, así como de Pedro Alzari y Héctor Viturro. A todos ellos les agradezco me hayan ayudado a superar un sinnúmero de dificultades.

Probablemente no estaría en condiciones de presentar esta tesis si no hubiera contado con la constante ayuda de Silvia Spinelli, que además de colaborar de muchas maneras, siempre realizó mis trámites. La edición y diagramación de la

misma es obra de Rita Bonetto y Carmen Paniagua, a quienes me resulta muy difícil expresarles el agradecimiento que siento por la paciencia y el cariño con que realizaron tan solidario esfuerzo.

Por último quiero agradecer:

- al personal no docente de la facultad de Ciencias Exactas.
- a mis amigos que me ayudan en todo.

