

Revista de la Facultad de
CIENCIAS de la UNI
REVCUNI

Revista de la Facultad de
CIENCIAS de la UNI
REVCUNI

Índice

Investigación

- Películas electrocrómicas de óxido de tungsteno y de óxido de hierro obtenidas por el proceso sol-gel.
H. Alarcón, Y. Zenitagoya, A. Medina y W. Estrada 03
- Estudio electrocrómico estructural y morfológico del óxido de níquel obtenido por rocío pirolítico.
M. Gómez, D. Rodríguez y W. Estrada 20
- El problema de la integral de Feynman.
H. G. Valqui 33
- El spin no es necesariamente relativístico.
H. G. Valqui 48

Tesis

- Diseño y construcción de un microscopio de efecto túnel.
Fernando Camino 59
- Estudio electrocrómico estructural y morfológico de películas delgadas de óxido de níquel preparadas por la técnica del rociado pirolítico.
Mónica Marcela Gómez León 75
- Análisis de excursiones de potencia por inserciones de reactividad en reactores de investigación.
Gerardo Lázaro Moreyra..... 81
- Síntesis y estudio de complejos de 3-(2-Piridil)-5,6-Difenil-1, 2, 4-Triazina con hierro (II) y cobre (I), su determinación espectrofotométrica y aplicación al análisis de agua pre-tratada de uso industrial.
Patricia Rosario López Pino 92
- Estudio de las propiedades estructurales ópticas y eléctricas de Sn O_x : F obtenido de soluciones pirolíticas con alto contenido de flúor.
Enrique Montoya Quiroz 102

Divulgación especializada

- Holografía.
Carmen Eyzaguirre Gorvenia, Aníbal Valera Palacios 113
- El vector de inercia.
H. G. Valqui 124
- Optimización usando el método del recocido simulado.
René Víctor Valqui Vidal 134
- Fe de errata: ¿Existe la longitud de una varilla? 143
- Fe de errata: El problema de los cuatro colores..... 143

Investigación

Películas electrocrómicas de óxido de tungsteno y de óxido de hierro obtenidas por el proceso sol-gel

H. Alarcón, Y. Zenitagoya, A. Medina y W. Estrada (*)

RESUMEN

Se obtuvieron películas electrocrómicas de óxido de tungsteno y óxido de hierro vía el proceso sol-gel, a partir de soluciones acuosas de $\text{Na}_2\text{WO}_4 \cdot 2\text{H}_2\text{O}$ (0,2M) y $\text{FeCl}_3 \cdot 6\text{H}_2\text{O}$ (0,37M) respectivamente. Las películas fueron depositadas sobre un sustrato de vidrio recubierto con ITO, mediante la técnica de recubrimiento por inmersión (dip-coating) a una velocidad de 0,05 cm/s, sometiéndolas luego a un tratamiento térmico a 500°C por 3 horas para la formación del óxido de tungsteno y 275° y 500°C por 30 minutos para el óxido de hierro. Las películas de óxido de tungsteno muestran una estructura

amorfa de WO_3 y las de óxido de hierro electrocrómico la fase $\gamma\text{-Fe}_2\text{O}_3$ (maghemita) y $\alpha\text{-Fe}_2\text{O}_3$ (hematita). Estudios con difracción de rayos X y transmitancia infrarroja para el WO_3 obtenida en forma de polvo indica que con tratamientos térmicos de 150°C se obtiene $\text{WO}_3 \cdot x\text{H}_2\text{O}$ y con tratamientos a 500°C se obtiene WO_3 monoclinico deshidratado.

El comportamiento electrocrómico fue estudiado en una celda electroquímica por voltametría cíclica usando como electrólito LiClO_4 (0,1 M) en carbonato de propileno para el óxido de tungsteno y NaOH (10^{-3} M) en agua para el óxido de hierro. La transmitancia óptica fue medida in-situ por un haz láser ($\lambda = 632,8 \text{ nm}$) obteniéndose un contraste óptico (del estado transparente al coloreado) del 50% para el óxido de tungsteno. La medida espectral de la transmitancia indica que el WO_3 presenta un electrocromismo pronunciado en todo el rango visible; en cambio el FeOx presenta electrocromismo pronunciado en las vecindades del cercano infrarrojo, a partir de 600 nm , con valores máximos del contraste óptico de 35%.

ABSTRACT

Tungsten oxide and iron based thin films were obtained by sol-gel route from starting water solutions of $\text{Na}_2\text{WO}_4 \cdot 2\text{H}_2\text{O}$ and $\text{FeCl}_3 \cdot 6\text{H}_2\text{O}$, respectively. The coatings were deposited on a glass substrate pre-coated by ITO; it was used the dipping process at $0,05 \text{ cm/s}$ followed by a heat treatment during 30 minutes at 500°C (tungsten oxide) and 500°C (iron oxide). The electrochromic tungsten oxide films showed an amorphous structure while the electrochromic iron oxide films showed the phase $\gamma\text{-Fe}_2\text{O}_3$. Electrochromism was studied in an electrochemical cell using cyclic voltammetry.

LiClO_4 (0,1 M) in PC and water solution of NaOH (10^{-3}) were used as electrolyte for tungsten oxide and iron oxide, respectively. Optical transmittance measurements in-situ ($\lambda = 632,8 \text{ nm}$) and ex-situ showed an optical contrast between the bleached and colored states of 50% for tungsten oxide and 35% for iron oxide.

Palabras claves:

Películas delgadas, sol-gel, electrocromismo, óxido de tungsteno, óxido de hierro.

Key Words:

Thin films, sol-gel, electrochromism, tungsten oxide, iron oxide.

Introducción y sumario

En el presente trabajo se realiza el estudio de las propiedades electrocrómicas que presentan las películas de óxido de tungsteno y del óxido de hierro con un enfoque hacia su aplicación en las llamadas "ventanas inteligentes". Se hace énfasis en la fabricación de las películas delgadas por la técnica sol-gel y su caracterización por modulación de su transmitancia óptica.

Los materiales electrocrómicos se caracterizan por presentar un cambio reversible y persistente de sus propiedades ópticas bajo la acción de un campo eléctrico aplicado; esto es, su transmitancia óptica puede variar en forma continua desde un valor máximo (estado transparente) hasta un valor mínimo (estado coloreado). Estos materiales tienen aplicación particularmente en las llamadas "ventanas inteligentes" las cuales controlan el paso de la energía radiante hacia los ambientes de una manera elegante y energéticamente eficiente; también son potencialmente útiles en espejos retrovisores con reflectancia variable y en contraste. Un estudio extenso sobre los materiales electrocrómicos se puede encontrar en las referencias [Fauhgnan et. al, 1995] y [Granqvist et. al., 1995].

Numerosas sustancias orgánicas e inorgánicas exhiben electrocromismo [Granqvist, 1995]. Entre éstas los óxidos de los metales de transición presentan particularmente gran interés por sus buenas cualidades electrocrómicas: contraste óptico pronunciado, reversibilidad, memoria, durabilidad, etc. El óxido de tungsteno hasta el presente es el material electrocrómico más estudiado [Granqvist, 1995], y hoy en día probablemente es el único que ha sido utilizado a escala industrial. Por otro lado, sorprendentemente, siendo el óxido de hierro uno de los más comunes dentro de la familia de los óxidos de metales de transición y de poseer varios estados de valencia, modificables mediante un tratamiento térmico o por un ciclaje voltamétrico en celdas electroquímicas, ha sido escasamente estudiado como material electrocrómico [Orel et. al., 1994]. Esto nos motivó a iniciar simultáneamente el estudio de ambos materiales utilizando un proceso de fabricación muy prometedor: la técnica del sol-gel.

El proceso sol-gel es una técnica sencilla y de bajo costo para la fabricación de películas delgadas y es potencialmente útil para aplicaciones en áreas extensas [Brinker et. al., 1990; Orel et. al; 1994]. La formación de las películas usando esta técnica es un proceso reproducible en donde las características estructurales y de composición de los materiales obtenidos pueden definirse a partir de las características físico-químicas de los materiales precursores, esto es de los soles y geles.

En general en la técnica sol-gel se parte de un alcóxido de un metal de transición y el proceso se realiza en atmósfera de humedad controlada. Este es el procedimiento más común para la fabricación de materiales electrocrómicos. En este trabajo se emplea un procedimiento alternativo más simple el cual no requiere de un sistema de humedad

controlada [Livage et. al, 1988; Chemseeddine, et. al; 1984]. Para el mismo se parte de una sal inorgánica disuelta en agua, cuyos componentes (denominados precursores) forman posteriormente soles homogéneos; por variación de su acidez, producto de una hidrólisis inicial seguida de un proceso de condensación, se produce una ulterior desestabilización de los soles dando origen a la formación de los geles. Dentro de los soles se sumerge el sustrato formándose una película interna que se adhiere al sustrato y una película externa que es devuelta al gel. Este mecanismo de recubrimiento es llamado "dip coating". Seguidamente se lleva a cabo un secado evaporando el solvente presente en la película; en la etapa final del secado, la película colapsa formándose una capa compacta, denominada xerogel. Finalmente se procede a un tratamiento térmico del xerogel dando origen a la película de óxido con diferentes grados de cristalinidad y composición química.

En este trabajo presentamos los primeros resultados obtenidos para los óxidos de tungsteno y de hierro producidos según el procedimiento sol-gel descrito. El comportamiento electrocrómico fue estudiado en una celda electroquímica de tres electrodos utilizando como electrolito LiClO_4 disuelto en carbonato de propileno para el óxido de tungsteno, y NaOH en agua para el óxido de hierro. La transmitancia óptica fue medida in-situ (en 632,8 nm) y ex-situ (espectralmente en el rango de 350 - 800 nm). Las características estructurales fueron estudiadas mediante la difracción de rayos x, transmitancia infrarroja y la espectroscopía Mossbauer por electrones de conversión (CEMS).

En la sección II se describe las condiciones experimentales en las que se realizó este trabajo. En la sección III se discute la preparación y la caracterización de las películas de óxido de tungsteno. Partiendo de una solución acuosa de $\text{Na}_2\text{WO}_4 \cdot 2\text{H}_2\text{O}$ (0,2M) y mediante un tratamiento térmico a 500°C por 3 horas el gel da origen a una película compacta de óxido de tungsteno. Mediante la difracción de rayos x se determinó que la estructura formada corresponde a WO_3 monoclinico.

Las películas presentan un electrocromismo pronunciado en la región espectral de 450-800 nm: la transmitancia varía alrededor del 50%, entre el estado transparente y el coloreado. En la sección IV se describe la preparación y la caracterización del óxido de hierro. Los recubrimientos fueron obtenidos a partir de una solución acuosa de $\text{FeCl}_3 \cdot 6\text{H}_2\text{O}$ (0,37 M); el gel fue sometido a tratamientos térmicos de 30 minutos, algunas a 275°C y otras a 500°C . Se obtuvo una variación en la transmitancia entre el estado transparente y coloreado de alrededor del 30% para el rango espectral entre 650-800 nm. La caracterización estructural de las películas de óxido de hierro se realizó por espectroscopía Mössbauer en la modalidad de transmisión (EMT). Finalmente en la sección V se resaltan las conclusiones más importantes.

Condiciones experimentales

Para la fabricación de las películas usando el método sol-gel se utilizaron reactivos químicamente puros cuyas soluciones acuosas se prepararon con agua tridestilada. Para el caso de los recubrimientos de óxido de tungsteno se partió de soluciones acuosas de $\text{Na}_2\text{WO}_3 \cdot 2\text{H}_2\text{O}$, y para el caso del óxido de hierro de soluciones acuosas de $\text{FeCl}_3 \cdot 6\text{H}_2\text{O}$. Los sustratos empleados fueron láminas de vidrio corning pre-cubiertos de una película transparente conductora eléctrica, In_2O_3 : Sn (ITO).

Para el análisis por difracción de rayos x (DRX) se utilizó un difractómetro Siemens con una fuente de $\text{Cu-K}\alpha$ y operado a 10 mA con un voltaje de 30 kV. Las medidas de transmitancia infrarroja (IR) se hicieron con un equipo Carl Zeiss Jena 73 (Specord). Para el análisis de espectroscopía Mossbauer se utilizó un equipo ELSCINT conectado a una PAC con tarjeta multicanal para la adquisición de los datos. Para realizar todas estas medidas no siempre se utilizó directamente las muestras película-sustrato, en su lugar se fabricó pastillas a partir del polvo del xerogel y del óxido.

Los experimentos de voltametría cíclica se realizaron con un potenciostato Wenking POS 73 conectado a una PC para la adquisición de los valores de voltaje y corriente. Se empleó una celda electroquímica de tres electrodos, en donde como electrodo de trabajo (ET) se utilizó a la muestra electrocrómica, como contraelectrodo o electrodo auxiliar (EA) se utilizó una lámina de platino y como electrodo de referencia Calomel Saturado (ER). Las medidas de voltametría cíclica se realizaron a una velocidad de barrido de 10 mV/s y fueron hechas "in-situ" con las medidas de transmitancia óptica monocromática (632,8 nm) para el cual se utilizó un haz de láser He-Ne y un fotodetector de silicio. Los valores de transmitancia que se reportan están referidos al sistema celda + electrólito + muestra, usándose como 100% de transmitancia a aquella que corresponde al sistema celda + electrólito. Las medidas de transmitancia espectral se realizaron con un espectrofotómetro OPTOMETRIC RS-325. Para el mismo se procedió de la siguiente manera: la muestra se sometió a voltametría cíclica en la celda electroquímica de tres electrodos; luego de un cierto número de ciclos, cuando la muestra adquiere su máxima transmitancia se la lleva al espectrofotómetro y se mide la transmitancia óptica. Luego se le retorna a la celda y se le aplica un potencial catódico hasta que alcance su mínima transmitancia la cual se mide nuevamente en el espectrofotómetro.

Fabricación y caracterización de películas delgadas de óxido de tungsteno

FABRICACIÓN

Los recubrimientos de óxido de tungsteno fueron producidos por el proceso sol-gel a partir de una solución acuosa de $\text{Na}_2\text{WO}_4 \cdot 2\text{H}_2\text{O}$. Todo el proceso experi-

mental seguido para la obtención de la película se representa esquemáticamente en la fig. 1.

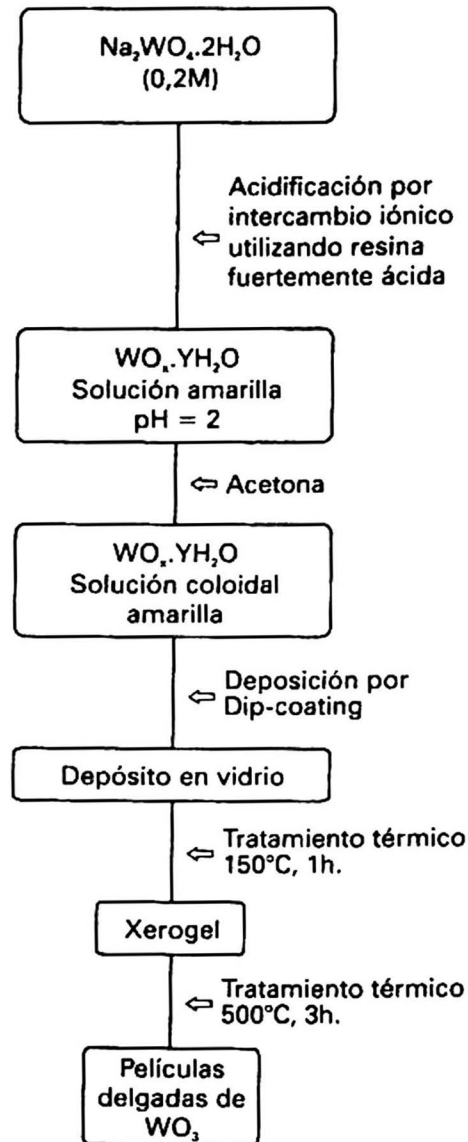


Fig. 1

Representación esquemática del proceso experimental seguido para la obtención de películas delgadas de óxido de tungsteno vía la técnica sol-gel.

Mediante un proceso de intercambio iónico, haciendo pasar la solución a través de una resina de intercambio catiónico fuertemente ácida, se obtienen especies condensadas distribuidas en un medio líquido de modo que el sistema globalmente podemos representarlo como $\text{WO}_3 \cdot x \text{H}_2\text{O}$, esto es, un óxido de tungsteno hidratado (ácido politungsténico); la concentración de las especies condensadas en la solución

gelificada depende de diversos parámetros tales como el pH, concentración de la sal, naturaleza del solvente, temperatura, etc. En el presente caso se obtuvo buenos resultados para soluciones gelificadas con $\text{pH} \sim 2$ y trabajando a la temperatura ambiente. El gel se estabiliza mediante la adición de un pequeño volumen de acetona ($\sim 10\%$ del volumen). Para encontrar el punto de trabajo, esto es las condiciones apropiadas para obtener un xerogel uniforme sobre el sustrato, se midió la variación de la transmitancia óptica monocromática del gel con el tiempo; se trabajó a $\lambda = 530$ nm y se usó diferentes concentraciones del tungstato de sodio. La fig. 2 muestra los resultados obtenidos para las concentraciones de 0,2 M; 0,3 M y 0,4 M; se encontró que para 0,3M y 0,4M las viscosidades eran excesivas a fin de permitir obtener un depósito uniforme del gel sobre el sustrato. A la concentración de 0,2M se encontró que valores de la transmitancia óptica mayores del 35% (viscosidad baja) no era posible lograr la adherencia de la solución sobre el sustrato; el rango óptimo para lograr una buena adherencia se determinó estar entre 22% - 27% de transmitancia, esto es, aproximadamente a los 30 minutos de permanecer en reposo la solución, después de haber finalizado el intercambio iónico por la resina. En este estado el sustrato es sumergido alternadamente en la solución coloidal a una velocidad de 0,05 cm/s. Luego de realizar 150 inmersiones, seguido de un tratamiento térmico a 500°C durante 3 h, se forma una película compacta de óxido de tungsteno sobre el sustrato. En adelante todos los resultados de la caracterización estructural y electrocrómica están referidos a la concentración de 0,2M.

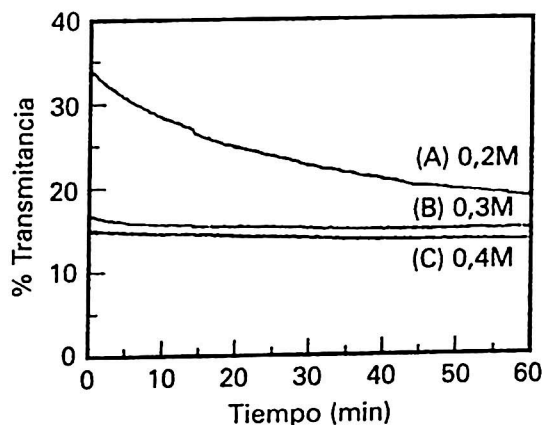


Fig. 2

Variación de la transmitancia óptica del ácido politungsténico a diferentes concentraciones de $\text{Na}_2\text{WO}_4 \cdot 2\text{H}_2\text{O}$ en función del tiempo.

CARACTERIZACIÓN

Las películas obtenidas según nuestra descripción previa fueron sometidas al análisis por difracción de rayos x. Los resultados indicaron que el WO_3 obtenido

tiene una estructura amorfa o eventualmente nano-cristalina ya que no se observó ningún pico de difracción. Sin embargo mediante un procedimiento alternativo se determinó la característica estructural que eventualmente la película adoptaría. En efecto, tomando porciones de 3-4 mg. del ácido politungsténico (gel) se procedió a someterlo a un tratamiento térmico de 500°C. Después de 3 horas de tratamiento térmico se formó un polvo conformado de estructuras cristalizadas; éstos fueron compactados en pastillas y sometidos al análisis por difracción de rayos x. La fig. 3 muestra el difractograma obtenido para muestras según las condiciones descritas. Todos los picos coinciden con el patrón de difracción del WO_3 monoclinico (según datos ASTM), siendo los más intensos los correspondientes a los ángulos 2θ : 23,1°; 23,7°; 24,3°; 33,3° y 34,1 atribuibiles a los planos cristalográficos (001), (020), (200), (021) y (201), respectivamente. Este resultado es coincidente con la estructura del WO_3 obtenido previamente en nuestro laboratorio usando la técnica del rociado pirolítico [Arakaki et. al., 1995].

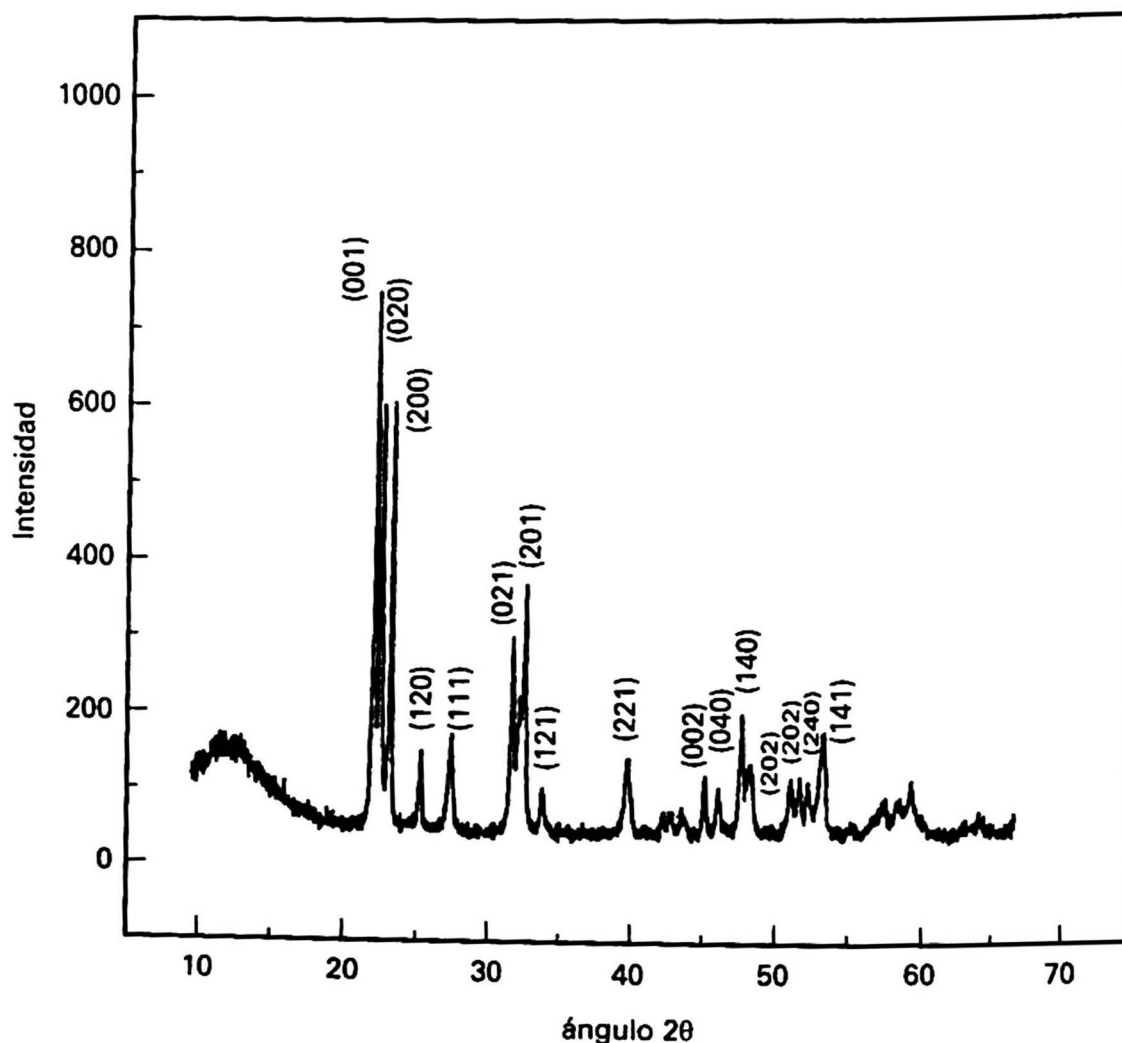


Fig. 3

Difractograma de rayos X del polvo de óxido de tungsteno obtenido siguiendo esquema de la fig. 1.

Adicionalmente se hicieron medidas de IR de los cristales obtenidos por tratamiento térmico del ácido politungsténico a las temperaturas de 150°C y 500°C, durante 1 y 3 horas respectivamente. La fig. 4 muestra estos resultados; la curva en la parte superior (fig. 4a) se obtiene para las muestras tratadas a 150°C. En ella se observa una banda ancha en la región 3300-3500 cm^{-1} y otra más aguda en la región 1500 - 1650 cm^{-1} . Estos picos son fácilmente atribuibles a los modos "stretching" y "bending" respectivamente del H_2O molecular dentro de la estructura del óxido. Así mismo se presenta una banda en 930 cm^{-1} , y en la región entre 530-800 cm^{-1} es notorio la superposición de dos bandas centradas aproximadamente en 650 y 670 cm^{-1} . El espectro inferior (fig. 4b) corresponde a una muestra tratada a 500°C; aquí se observa picos solo en la región 500-1000 cm^{-1} ; estos están superpuestos y centrados aproximadamente en 750 cm^{-1} y 810 cm^{-1} . Estos no son atribuibles a alguno de los modos de vibración del H_2O . Estos resultados pueden interpretarse a partir de los trabajos de [Burdís et. al, 1994], [Chemssedine et. al, 1984] y [Taylor et. al, 1994].

En efecto, el pico en 930 cm^{-1} , en la muestra tratada térmicamente a 150°C (fig. 4a), es atribuible a los enlaces terminales $\text{W}=\text{O}$ producto de los oxígenos producidos por la incorporación de H_2O en la estructura octaédrica del óxido de tungsteno. Este pico no se observa para la muestra tratada a 500°C (fig. 4b), confirmando la deshidratación de la muestra a este nivel de temperatura. La banda en la región entre 530-800 cm^{-1} (fig. 4a), atribuible a las vibraciones $\text{W}-\text{O}-\text{W}$, se corre hacia regiones de frecuencias mayores en las muestras tratadas a 500°C (fig. 4b), observándose además que el desdoblamiento de la banda en dos picos es claramente más pronunciado para este último caso.

Este resultado puede ser interpretado como que existe una ligera distorsión de la unidad básica octaédrica del WO_3 en las muestras tratadas a 150°C, debido a la incorporación de H_2O indicando que a este nivel del tratamiento térmico no es aún suficiente producir una deshidratación completa de la muestra. De estos resultados se desprende que la estructura xerogel constituye esencialmente una estructura WO_3 (cuasimonoclínica) hidratada en donde los niveles de hidratación son determinados por la temperatura del tratamiento térmico. La estructura compacta, como película delgada, va adoptando la estructura WO_3 monoclínica a medida que aumentan los niveles de deshidratación de la muestra.

Para caracterizar el comportamiento electrocrómico de las películas de óxido de tungsteno, éstas fueron sujetas a un tratamiento electroquímico en una celda de tres electrodos conteniendo una solución de LiClO_4 (0,1M) en carbonato de polipropileno y se procedió según se describió en la sección II. Los recubrimientos obtenidos con tratamientos térmicos relativamente inferiores a 500°C presentan electrocromismo pero se degradan muy rápidamente haciéndolos aún inadecuados para aplicaciones

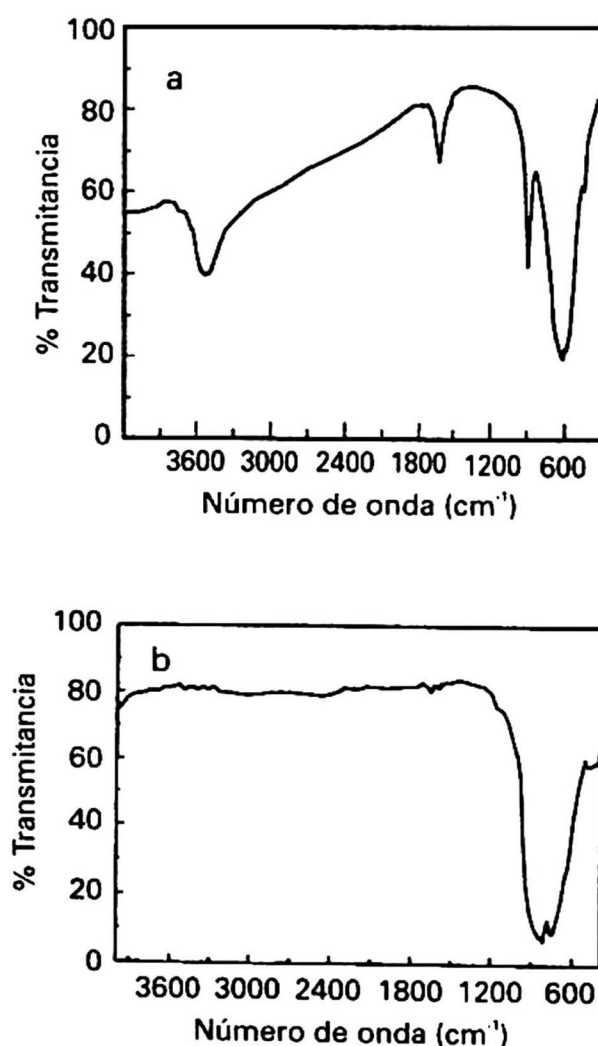
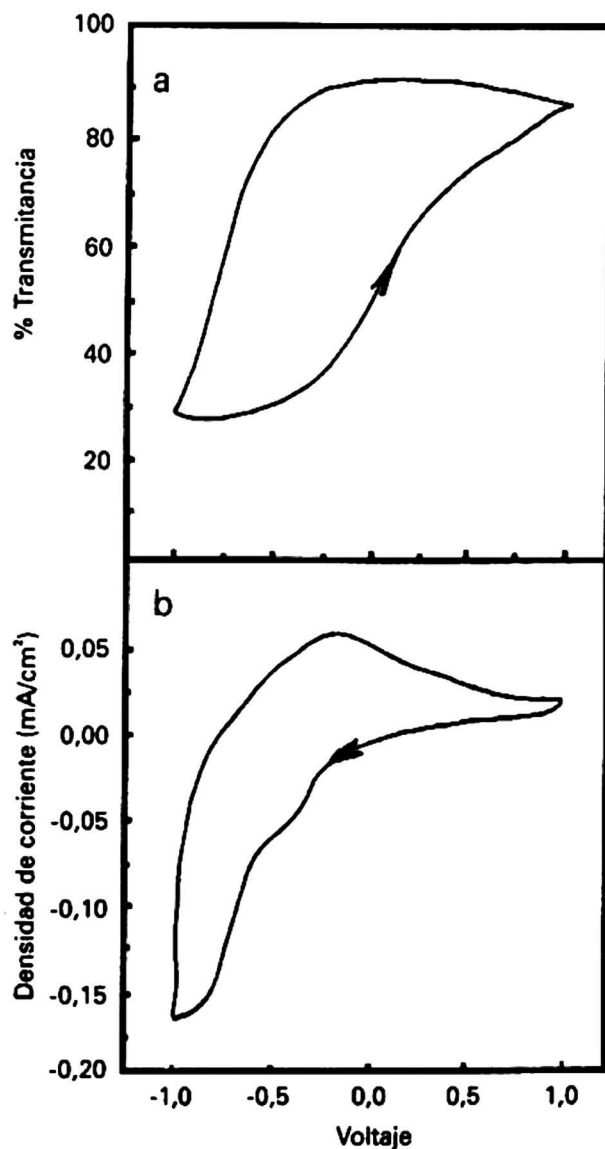


Fig. 4(a) y 4 (b)

Transmitancia Infrarroja del (a) xerogel, obtenido con un tratamiento térmico a 150°C, 1h, (b) óxido de tungsteno, obtenido a 500°C, 3h.

electrocrómicas. La fig. 5 muestra la voltametría cíclica in-situ con la medida de la transmitancia óptica ($\lambda = 632,8$ nm), después de cinco ciclos voltamétricos, para una muestra tratada térmicamente a la temperatura de 500°C durante 3 horas. Aquí se puede observar un pico anódico "ancho" alrededor de -0,5 V el cual correlacionándolo con las medidas in-situ de la transmitancia se deduce que la inserción del catión ocurre durante el proceso de coloración. Así mismo se observa que a este número de voltametrías cíclicas se obtiene un contraste óptico entre el estado transparente y el estado coloreado de ~50%.

**Fig. 5**

Transmitancia óptica en $\lambda = 632,8 \text{ nm}$ in-situ (a) y voltograma cíclico (b) del óxido de tungsteno obtenido por el proceso sol-gel, usando vidrio recubierto de ITO como sustrato. La película de WO_3 fue obtenida con 150 inmersiones seguido de un tratamiento térmico a 500°C durante 3 horas.

La fig. 6 muestra la transmitancia espectral del estado transparente y coloreado en el rango visible entre 450 y 800 nm, para la película correspondiente a la fig. 5,

después de 15 ciclos voltamétricos. Se observa que para 630 nm hay un contraste óptico de ~50% indicando que el material electrocrómico ha alcanzado su estado "estable". Este comportamiento se mantiene a medida que aumenta el tratamiento electroquímico; así mismo se observó que después de 20 ciclos aparecen ligeros signos de degradación mecánica en la película, los cuales son claramente notorios después de 50 ciclos.

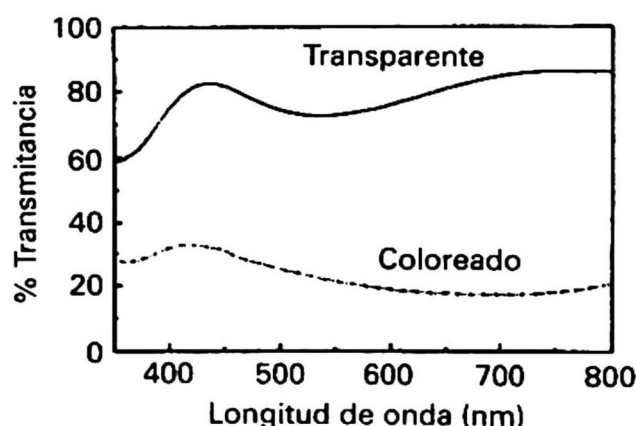


Fig. 6

Transmitancia espectral de la misma película de óxido de tungsteno mostrada en la fig. 5.

A estas alturas todavía no se tiene una completa explicación del origen de este proceso de degradación, pero presumimos que la adherencia del recubrimiento sobre el sustrato está jugando un papel importante. Actualmente estamos ejecutando nuevos diseños experimentales para mejorar la durabilidad de las muestras, los mismos que serán reportados próximamente.

Fabricación y caracterización de películas delgadas de óxido de hierro

FABRICACIÓN

Las películas de óxido de hierro fueron obtenidas por la técnica sol-gel a partir de una solución acuosa de $\text{FeCl}_3 \cdot 6\text{H}_2\text{O}$ (0,37 M). Se adiciona a ésta NH_3 (ac) al 25%, lentamente, hasta que alcance la solución un pH ~ 9-10. Se forma un precipitado marrón, el cual se centrifuga y se lava con agua tridestilada con el fin de eliminar el exceso de iones NH_4^+ (ac) y Cl^- (ac). La polimerización se lleva a cabo adicionando lentamente CH_3COOH (ac) al 60%, dando lugar a la formación de los soles de hierro después de 12 horas. Después de filtrar el sol formado, se ajusta su viscosidad con

glicerol. Finalmente, el sustrato (vidrio pre-recubierto con ITO) es sumergido (y extraído) en el (del) sol viscoso a una velocidad de 0.05 cm/s; luego de 30 inmersiones se forma sobre el sustrato una capa líquida el cual luego de un tratamiento térmico durante 30 minutos da origen a una película compacta de óxido de hierro sobre el sustrato. Se emplearon dos temperaturas típicas para el tratamiento térmico: 275°C y 500°C. La fig. 7 muestra de manera esquemática el procedimiento experimental seguido.



Fig. 7

Representación esquemática del proceso experimental seguido para la obtención de películas delgadas de óxido de hierro vía la técnica sol-gel.

CARACTERIZACIÓN

La fig. 8 muestra los espectros Mössbauer obtenidos para la muestra de óxido de hierro; aquí la línea continua es el resultado de la curva de ajuste y cuyos parámetros

hiperfinos se especifican en la tabla 1. La fig. 8a presenta el espectro CEMS para la muestra tratada a 275°C. Aquí doblete presente puede ser asociado a estados de Fe^{3+} y por los valores de IS (corrimiento isomérico) y QS (momento cuadrupolar), según se indica en la tabla 1, se puede deducir [Stevens et. al, 1975] que a este nivel de tratamiento térmico en la película hay una predominancia de la fase $\gamma - \text{Fe}_2\text{O}_3$ (maghemita).

Tabla 1
Parámetro Mössbauer de la muestra tratada a 275°C.

IS (mm/s)	QS (mm/s)	Γ (mm/s)	H (kG)
0,37 (5)	0,86 (1)	0,52 (1)	-

Tabla 1
Parámetro Mössbauer de la muestra tratada a 275°C.

IS (mm/s)	QS (mm/s)	Γ (mm/s)	H (kG)
0,38 (2)	-0,220 (1)	526	16,9

En la fig. 8 (b) se muestra el espectro Mössbauer de la muestra tratada a 500°C. De aquí se deduce la existencia de un doblete que puede ser asociado a estados de Fe^{3+} ; el sexteto presente posee un campo hiperfino (BHF) y un corrimiento isomérico (ver tabla 1) que pueden ser asociados a la fase $\alpha - \text{Fe}_2\text{O}_3$ (hematita), aunque el valor del desdoblamiento cuadrupolar es algo diferente a la de la hematita. Esto estaría indicando que existe una ligera distorsión del entorno electrónico de los átomos de hierro en las películas comparadas con el entorno electrónico del hierro en la hematita.

De estos resultados se desprende que las fases de óxido de hierro que se pueden finalmente obtener como película delgada dependen del nivel de tratamiento térmico. Actualmente estamos estudiando la influencia de las características del sol sobre las propiedades estructurales y morfológicas de la película. Para ello estamos caracterizando soles mediante microscopía electrónica y métodos de dispersión de luz ("light scattering"). Estos resultados serán publicados próximamente.

El electrocromismo fue estudiado utilizando una celda electroquímica de tres electrodos, usando como electrolito una solución acuosa de NaOH (10^{-3} M). Se siguió un procedimiento análogo al descrito para el óxido de tungsteno. De las medidas "in situ" se dedujo que, contrariamente al WO_3 , para el óxido de fierro la coloración es anódica esto es que el estado de coloración más pronunciado se produce por extracción de cationes. La fase $\gamma - \text{Fe}_2\text{O}_3$ mostró un electrocromismo ligero, con una variación de la transmitancia $< 15\%$, y la fase $\alpha - \text{Fe}_2\text{O}_3$ no mostró ningún cambio de su transmitancia óptica monocromática.

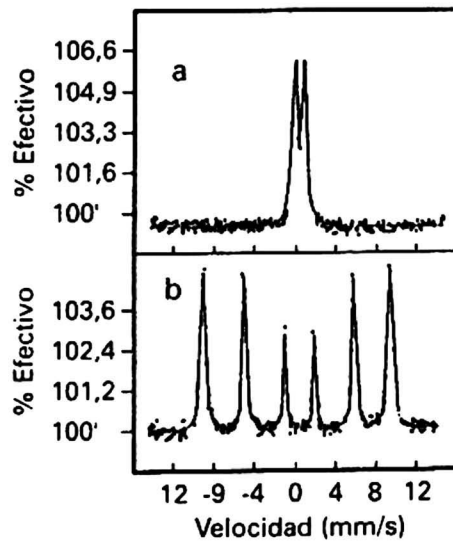


Fig. 8

Espectros CEMS para las películas de óxido de hierro tratadas térmicamente a (a) 275 y (b) 500°C.

Sin embargo las medidas ex-situ de transmitancia espectral indican un resultado interesante en el comportamiento electrocrómico del FeO_x . En efecto los resultados de la transmitancia espectral indica que no hay fenómeno electrocrómico en el óxido de hierro en su fase α para todo el rango visible; sin embargo este fenómeno para la fase γ se manifiesta a partir de los 530 nm. La fig. 9 muestra la variación de la transmitancia para el $\gamma - \text{Fe}_2\text{O}_3$ de aquí podemos observar que el efecto electrocrómico se va acentuando a medida que nos acercamos a la zona del cercano infrarrojo.

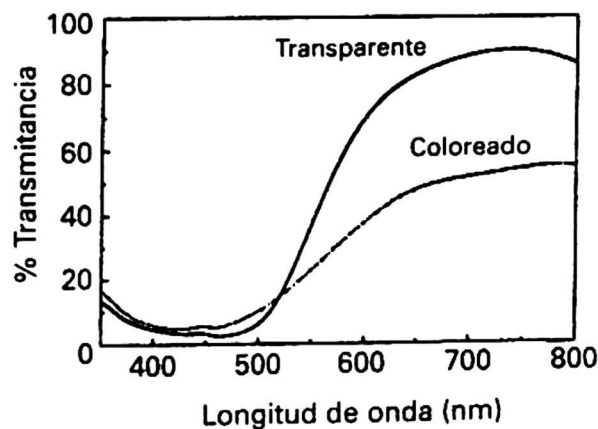


Fig. 9

Transmitancia espectral de la película de óxido de hierro obtenida por el proceso sol-gel.

Así, para 700 nm, por ejemplo, podemos deducir un contraste óptico entre el estado transparente (curva continua) y el estado coloreado (curva entrecortada) de 35%.

Aún no tenemos una explicación completa de las causas de este comportamiento electrocrómico diferenciado entre las fases del óxido de hierro y estamos actualmente dando pasos tanto teóricos como experimentales para comprender este interesante efecto.

Conclusiones

Usando una variante sencilla sol-gel, sin el requerimiento de atmósfera de humedad controlada, hemos obtenido películas de WO_3 y FeO_x con buenas características electrocrómicas. La estructura y morfología de estos recubrimientos son fuertemente dependientes del nivel de temperatura del tratamiento térmico; así mismo las características electrocrómicas son muy sensibles a estas características estructurales. Los recubrimientos de WO_3 compactadas a temperaturas relativamente menores a 400°C se degradan rápidamente; sobre esta temperatura el electrocromismo mejora sustancialmente tanto en contraste óptico como en durabilidad. Los recubrimientos de FeO_x presentan electrocromismo sólo en su fase γ , siendo particularmente notable el hecho que este fenómeno se presenta en regiones vecinas al cercano infrarrojo. La fase α no presenta electrocromismo.

REFERENCIAS

- Arakaki, J.; R. Reyes; M. Horn and W. Estrada. *Solar Energy Matter*, 37 (1995) 33.
- Brinker C. J. and G. W. Scherer. *Sol-Gel Science*, Academic Press, London (1990).
- Burdiss M. and J. Siddle, *SPIE*, Vol. 2255 / 371 (1994).
- Chemseeddine A., M. Henry and J. Livage, *Revue de Chimie Minérale*, T. 21, pag. 487, (1984).
- Fauhgnan B. W. and R. S. Crandall in *Display Devices*, Ed. by J. L. Pankove, *Topics in Advanced Physics*, Vol. 40, pag. 181, Springer, Berlin (1980).
- Granqvist C. G., *Handbook of Inorganic Electrochromic Materials*, Elsevier Science, 1995.
- Livage J., M. Henry and C. Sánchez, in *Prog. Solid. St. Chem.*, Vol. 18, pag. 259. Pergamon Press Plenum, Great Britain (1988).
- Orel B., M. Macek, and F. Sveg, *Thin Solid Films*, 246, 131 (1994).
- Stevens J. G. and V. E. Stevens. *Mossbauer effect data index*. IFI/PLENUM, New York (1975).
- Taylor T. A. and H. H. Patterson, *Appl. Spectroscopy*, Vol. 48, N° 6 (1994).

Agradecimientos

Agradecemos por la ayuda prestada en la toma de los espectros Mossbauer de las películas de óxido de hierro, al Dr. Juan Dávalos miembro del instituto de física Rocasolano de Madrid, España.

De la misma forma agradecemos al Mg. José Solís por la toma del espectro de rayos X.

Igualmente agradecemos el apoyo prestado en la toma de los espectros infrarrojo a la Lic. Esther Ocola docente de la Facultad de Ciencias de la Universidad Nacional de Ingeniería.

Agradecemos así mismo al International Program for Physics Science y a CONCYTEC por la ayuda económica prestada.

Agradecemos al Instituto de Investigación de la Facultad de Ciencias de la UNI, donde se elaboró el presente trabajo.

Estudio electrocrómico, estructural y morfológico del óxido de níquel obtenido por rocío pirolítico

M. Gómez, D. Rodríguez y W. Estrada (*)

RESUMEN

Usando la técnica del rocío pirolítico (spray pyrolysis) se obtuvo el óxido de níquel electrocrómico en forma de película delgada. Los recubrimientos presentan una transmitancia óptica en el estado claro de $\approx 85\%$ (para $\lambda = 630 \text{ nm}$), con una modulación óptica entre el estado claro y el estado coloreado de $\approx 50\%$.

Las películas fueron fabricadas siguiendo dos procedimientos: (a) usando como solución pirolítica nitrato de níquel disuelto en alcohol, (b) agregando pequeños porcentajes de nitrato de cobalto a las soluciones acuosas de nitrato de níquel. Los recubrimientos obtenidos mediante (a) muestran una estructura de NiO-bunsenita, mientras que aquellos obtenidos mediante (b) presen-

(*) Facultad de Ciencias, Universidad Nacional de Ingeniería, Casilla postal 31-139. Lima.

tan una mezcla de fases de NiO y Co₃O₄. Ambos procedimientos producen recubrimientos con propiedades electrocrómicas comparables a aquellas obtenidas por técnicas de vacío: buen contraste óptico, reversibilidad y durabilidad.

ABSTRACT

Thin films of nickel-oxide-based coatings have been obtained using the spray pyrolysis technique. The coatings show electrochromism with an optical transmittance of $\approx 85\%$ (at $\lambda = 630$ nm) in the bleached state, and an optical modulation of $\approx 50\%$ between the bleached and colored states.

The coatings were fabricated following two procedures: (a) using a pyrolytic solution of nickel nitrate dissolved in alcohol, (b) mixing cobalt nitrate and nickel nitrate in water solution. Films obtained by procedure (a) show a NiO-bunsenite structure while those obtained by procedure (b) present the NiO and Co₃O₄ phases. These two procedures give coatings with electrochromic properties as good as those obtained by vacuum techniques: high optical contrast, reversibility and durability.

Palabras claves

Spray pyrolysis	Rocío pirolítico
Electrochromism	Electrocromismo
Thin films	Películas delgadas
Nickel oxide	Oxido de níquel
Cobalt oxide	Oxido de cobalto

Introducción

El electrocromismo es un fenómeno que despierta mucho interés por sus potenciales usos en las llamadas "ventanas inteligentes" (smart windows), el cual permite balancear la energía térmica y la luminosidad de los ambientes, en los espejos retrovisores de autos, en las pantallas de información no emisivas, etc. Un material electrocrómico se caracteriza por cambiar en forma reversible y persistente sus propiedades ópticas cuando es sometido a la acción de un campo eléctrico. Un dispositivo electrocrómico es un sistema que básicamente está compuesto por una capa electrocrómica, un contraelectrodo y un medio conductor iónico. La modulación óptica está asociada a la inserción y extracción de cationes al material electrocrómico por medios electroquímicos. Excelentes tratados sobre los materiales electrocrómicos así como de dispositivos se dan en las siguientes referencias: Lampert and Graqvist (1988), Carpenter and Corrigan (1990), and Granqvist (1995).

Uno de los materiales que ha recibido mucha atención es el óxido de níquel debido a sus buenas propiedades electrocrómicas: contraste óptico pronunciado, durable, tiempo de respuesta adecuado para los "smart windows", etc. Para la obtención del óxido de níquel en forma de película delgada con propiedades electrocrómicas se han empleado varios tipos de técnicas, destacándose que mayoritariamente predominan aquellas basadas en métodos de vacío [Svensson and Granqvist, 1987, Estrada et al, 1988] y aquellas basadas en técnicas electroquímicas [Fantini and Goresntein, 1987, Carpenter et al, 1987]. Recubrimientos electrocrómicos usando el método del rocío pirolítico ha sido escasamente empleado [Arakaki et. al, 1995], a pesar de ser una técnica de bajo costo, sencilla y potencialmente útil para recubrimientos en áreas grandes.

En un trabajo previo [Arakaki et. al, 1995] publicamos el comportamiento electrocrómico del NiO_x obtenido por rocío pirolítico a partir de una solución acuosa de nitrato de níquel. Estos recubrimientos mostraron un marcado electrocromismo, aunque con una transmitancia óptica para el estado transparente relativamente baja si se piensa en aplicaciones prácticas ($\approx 60\%$). Este efecto se asoció principalmente al carácter rugoso de la superficie del recubrimiento, intrínseco a la interacción del sustrato caliente con las soluciones acuosas del nitrato de níquel.

En este trabajo se presentan superficies de NiO_x substancialmente mejoradas usando la misma técnica del rocío pirolítico; las películas presentan un electrocromismo pronunciado y con alta transmitancia en el estado claro. Para este resultado se realizaron dos procedimientos que dieron igualmente resultados satisfactorios: (a) empleando una solución alcohólica de nitrato de níquel, (b) usando una solución acuosa de la mezcla de nitrato de níquel con nitrato de cobalto.

Experimento

Las películas de NiO_x fueron obtenidas mediante un sistema de rocío similar al descrito en un trabajo previo [Arakaki et. al. 1995], utilizando aire filtrado como gas portador del aerosol. El electrocromismo se estudió mediante medidas in-situ y ex-situ de la voltametría cíclica con la transmitancia óptica. Para la voltametría cíclica (inserción/extracción de carga) se utilizó un potentiostato WENKING POS 73 en una celda electroquímica de tres electrodos: la muestra como electrodo de trabajo (ET), una lámina de platino como contraelectrodo (CE) y calomel saturado como electrodo de referencia (ER). La transmitancia óptica fue medida con un sistema láser He-Ne -detector- registrador y un espectrofotómetro Optometric RS 375 para el caso monocromático y espectral, respectivamente. El proceso de inserción/extracción de carga se realizó en un medio electrolítico que consistió de una solución acuosa básica de KOH 0,1M.

Como sustrato se empleó portaobjetos de vidrio pre-cubiertos de una película transparente y conductora eléctrica, comercialmente conocido como ITO. El espesor

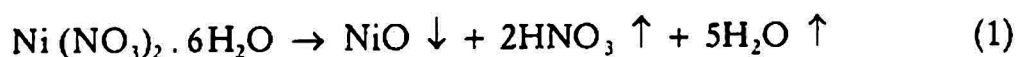
de los recubrimientos fueron medidos por el método de la interferometría óptica. Las soluciones pirolíticas empleadas en los procedimientos mencionados fueron las siguientes: soluciones alcohólicas de nitrato de níquel con concentraciones de 0,05; 0,1 y 0,25 M; soluciones acuosas de la mezcla de nitrato de níquel: nitrato de cobalto en las proporciones molares de 100:0, 90:10, 70:30 y 50:50.

Para el análisis estructural de las películas se empleó difracción de rayos X, y para el estudio de la relación Ni: Co en las películas se empleó fluorescencia de rayos X.

Resultados y discusión

PELÍCULAS OBTENIDAS A PARTIR DE SOLUCIONES ALCOHÓLICAS DE NITRATO DE NÍQUEL

Las películas obtenidas a partir de soluciones alcohólicas fueron crecidas a la temperatura de 300 °C. A este nivel de temperatura se presume que el solvente alcohólico no interviene en la reacción pirolítica, siendo su papel básicamente como transportador del aerosol. Considerando que la sal empleada era hexahidratada probablemente la reacción química que se produce en el sustrato, dando origen a la fase sólida de óxido de níquel, pueda globalmente ser representada por:



Se encontró que las velocidades de crecimiento de los recubrimientos tenían una dependencia directa de la concentración de nitrato en la solución. Desde que el límite de solubilidad del nitrato de níquel en el alcohol es de 0,3 M; los mejores resultados (compatibilizando el espesor con la adherencia) se obtuvieron para las concentraciones de 0,25 M. Por consiguiente los parámetros óptimos experimentales se encontraron a T=300 °C y concentración de 0,25 M. Los resultados de caracterización que se indican en lo que sigue son básicamente referidos a estas condiciones de preparación.

La fig. 1(a) muestra el resultado del patrón de difracción de una película de óxido níquel crecido sobre un sustrato de vidrio pre-cubierto con ITO. El espesor de la película fue de ~500 nm para un tiempo de rocío de 30 minutos. De ella puede deducirse que la estructura corresponde al NiO cúbico del tipo Bunsenite. En la fig. 1 (b) se muestra el patrón de difracción de rayos X para la misma película después de haber sido sometido sucesivamente a 3000 ciclos voltamétricos en un sistema electroquímico tal al descrito en la sección anterior. Aquí puede observarse claramente que para este último caso todos los picos presentan perfiles más agudos respecto a las muestras no-cicladas, pero particularmente se observa que este efecto es notoriamente más pronunciado en la dirección (200). Aún no tenemos una explicación completa

de este resultado, pero el mismo podría ser atribuido a los siguientes factores: recristalización de las regiones amorfas de la estructura durante la inserción/extracción de cationes, particularmente en los bordes de grano en donde hay alta concentración de defectos; otra posibilidad podría ser el crecimiento ulterior de nuevas capas de NiO durante el ciclaje electroquímico. Recientes estudios que estamos realizando en el laboratorio mediante microscopía electrónica indican un claro proceso de envejecimiento en los bordes de grano compatibilizando mejor con la primera hipótesis. Estos resultados y mayores elementos que nos permitan entender esta "recristalización" se viene estudiando y será materia de una próxima publicación.

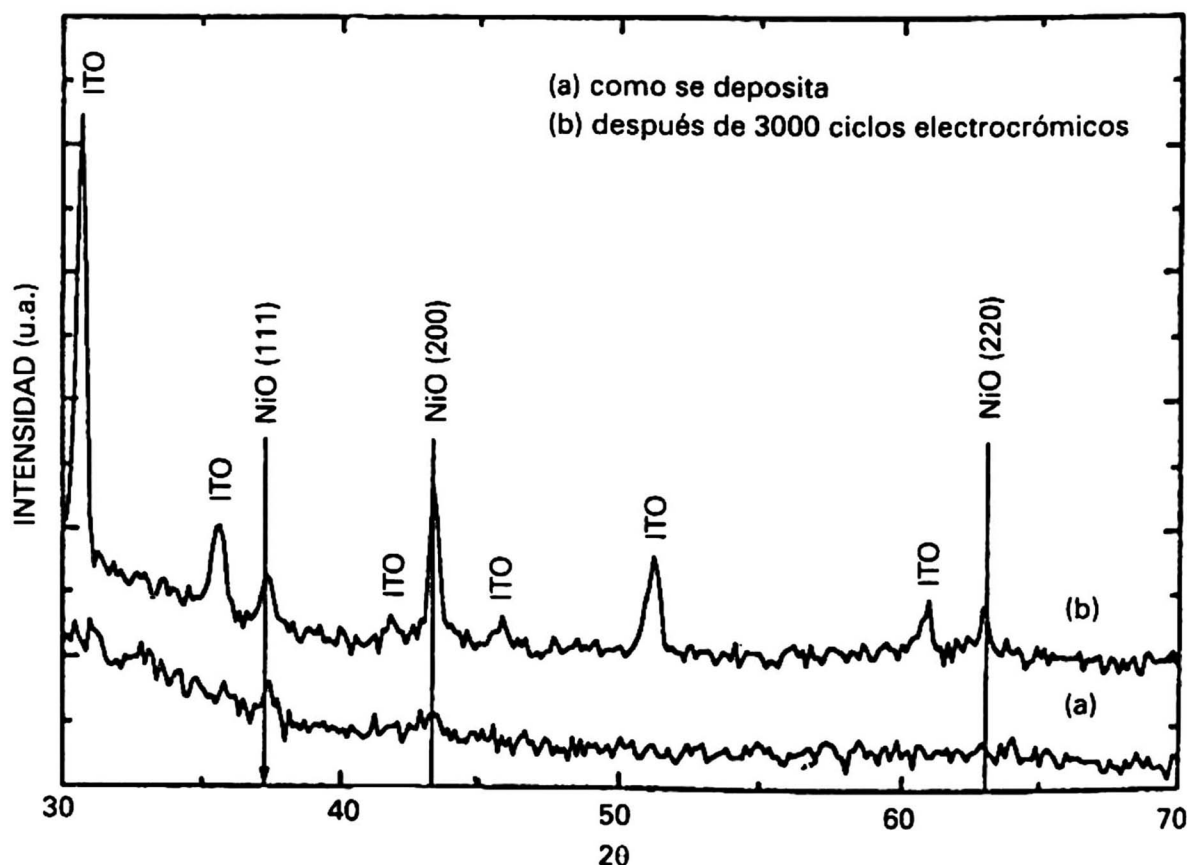


Fig. 1
Difractogramas de rayos X de películas delgadas de óxido de níquel preparadas empleando una solución 0,25 M de $\text{Ni}(\text{NO}_3)_2 \cdot 6\text{H}_2\text{O}$ a 300°C . (a) Corresponde a la película tal como se deposita sobre vidrio, y (b) a la película depositada sobre ITO después de ser sometida a 3000 ciclos voltamétricos.

El electrocromismo del NiO_x evoluciona paulatinamente conforme se va aumentando el número de ciclos voltamétricos. En la fig. 2 se muestra la transmitancia

óptica monocromática (parte superior) in-situ con la voltametría cíclica (parte inferior). De aquí se puede deducir varios detalles. En primer lugar que la cantidad de carga insertada/extraída (área positiva/negativa del perfil voltamétrico con el eje del potencial) crece paulatinamente a medida que aumenta el número de ciclos.

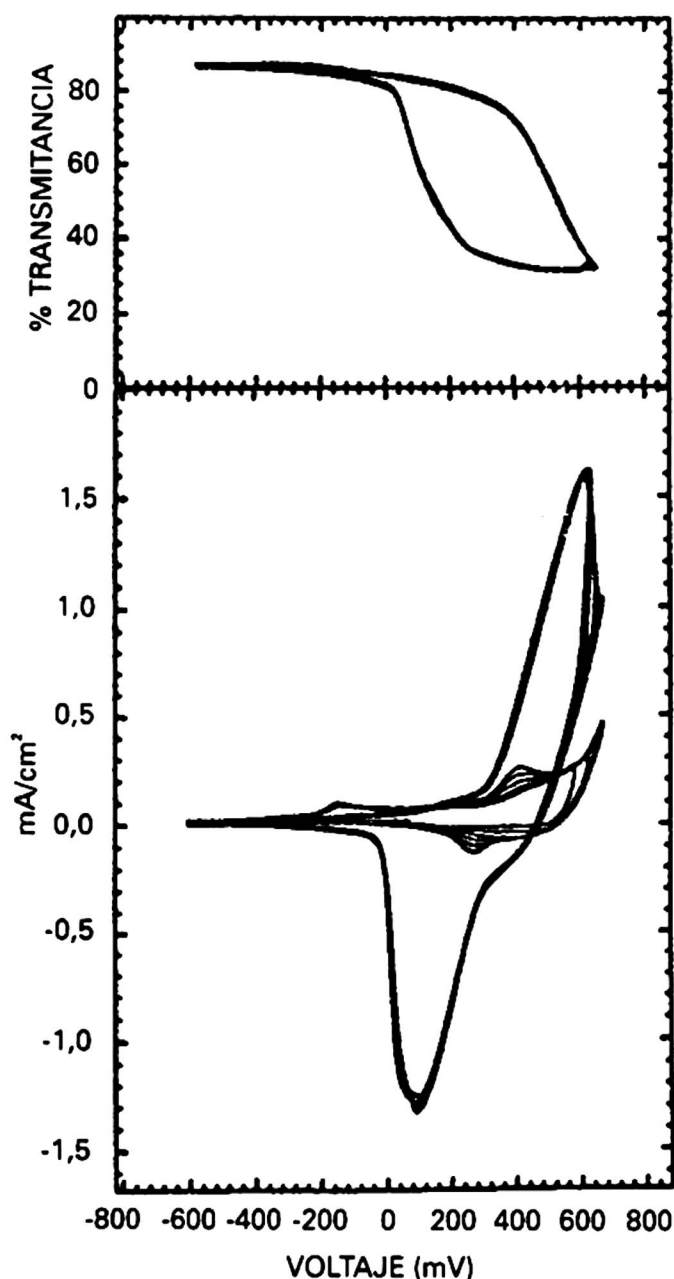


Fig. 2

Voltametría cíclica (parte inferior) asociada a la transmitancia in-situ (parte superior), de una película delgada de óxido de níquel preparada por rocío pirolítico partiendo de una solución alcohólica de $\text{Ni}(\text{NO}_3)_2 \cdot 6\text{H}_2\text{O}$ 0,25 M y depositada a 300°C.

En efecto para los primeros ciclos (perfiles pequeños en la parte inferior de la fig. 2) se tiene un valor aproximado de carga insertada/extraída de 6 mC/cm^2 ; luego de 20 voltametrías cíclicas el perfil voltamétrico crece hasta corresponder a una cantidad de carga insertada extraída de 30 mC/cm^2 . Para ulteriores incrementos de ciclos voltamétricos el perfil no experimenta mayor variación. El perfil mayor mostrado en la fig. 2 (parte inferior) se mantiene para más de 3000 ciclos. A este estado límite se le denomina **estado electrocrómico "estable"**, en donde se produce un proceso reversible de inserción y extracción de carga. La curva de transmitancia (parte superior de la fig. 2) corresponde al estado "estable". De ella podemos observar que la transmitancia óptica varía desde 85% de transmitancia hasta 35% conforme se va insertando carga al ET, indicando un proceso de coloración anódico.

En la fig. 3 se muestra la transmitancia espectral para la misma película en sus estados de mayor transparencia y de mayor coloración. Se observa que el electrocromismo es pronunciado en la región de 400-600 nm. Por ejemplo para 550 nm de longitud de onda se alcanza una variación entre los estados claro y coloreado de $\sim 50\%$. Este resultado indica claramente que las películas electrocrómicas de NiO_x obtenidas por rocío pirolítico a partir de soluciones alcohólicas poseen características electrocrómicas tan competentes como aquellas producidas por "sputtering" [Estrada et. al, 1988] con el adicional de ser bastante más sencillas y de menor costo.

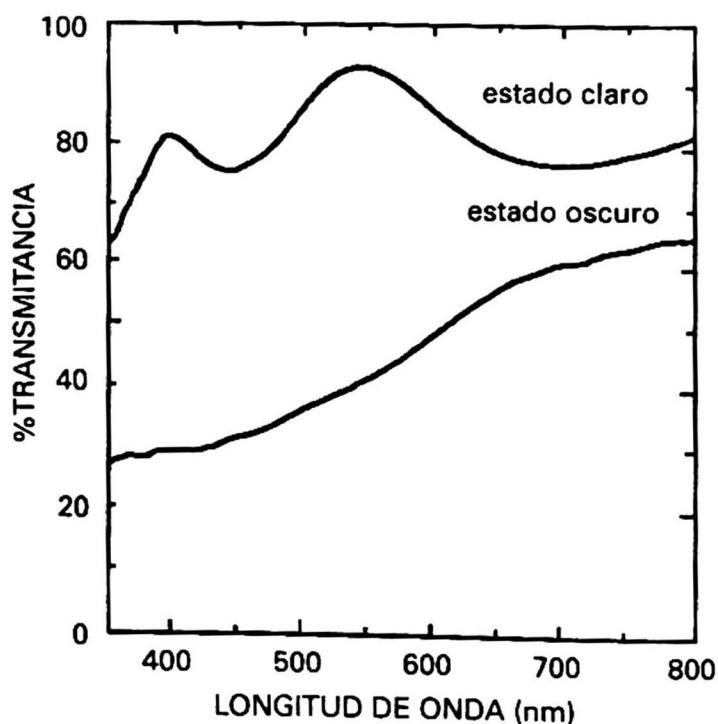


Fig. 3

Transmitancia espectral óptica de una película delgada de óxido de níquel en sus estados claro y oscuro. Las condiciones de preparación son las mismas que las de la figura 2.

PELÍCULAS OBTENIDAS A PARTIR DE SOLUCIONES ACUOSAS DE NITRATO DE NI: CO

Un proceso alternativo a fin de obtener NiO_x con propiedades electrocrómicas óptimas se obtuvo a partir de soluciones provenientes de mezclas acuosas de nitrato de níquel y nitrato de cobalto. Para la obtención de las películas y el análisis electrocrómico ulterior se procedió de una manera similar al descrito en la parte (a). Utilizando diferentes proporciones molares de nitrato de níquel: nitrato de cobalto según especificadas en la sección anterior se procedió a rociar la solución sobre el sustrato caliente a 300°C . La fig. 4 muestra el contenido de Co (% atómico) en la película versus el contenido de cobalto en la solución. Estos datos fueron extraídos mediante medidas de fluorescencia de rayos X. De este resultado podemos entonces establecer una correlación directa entre el contenido de nitrato de cobalto en la solución con el % atómico de Co en la película.

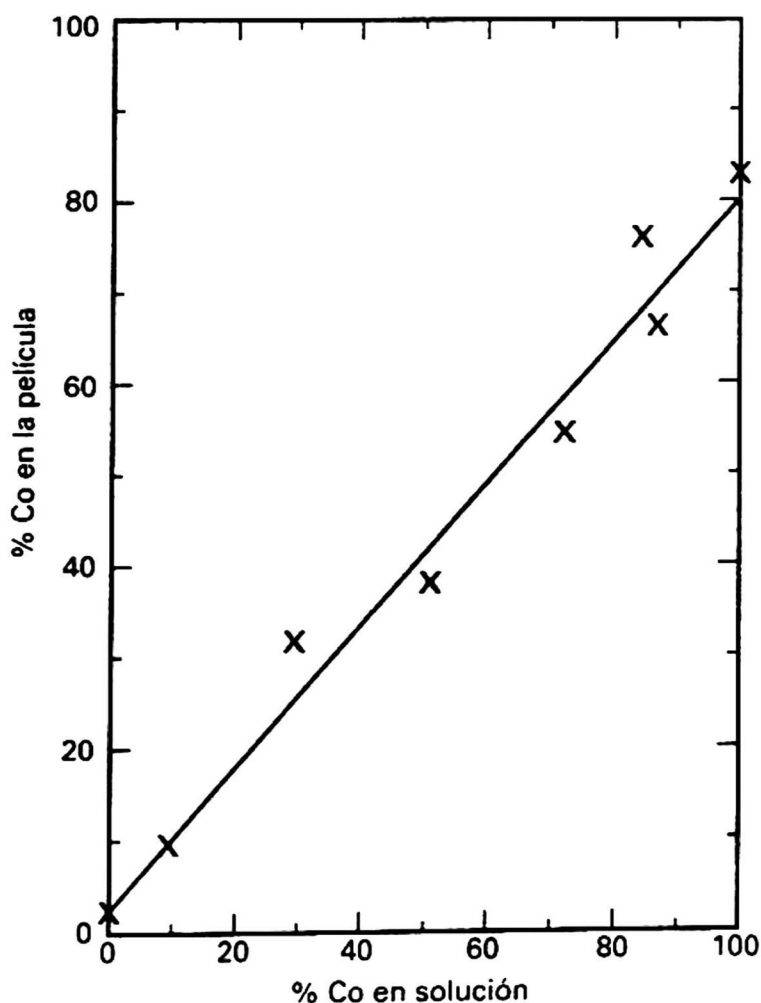


Fig. 4

Medidas de fluorescencia de rayos X que relaciona los contenidos de Co (% atómico) en la película delgada con la concentración de Co en la solución que se emplea para su depósito.

Las diversas caracterizaciones electrocrómicas de las películas obtenidas indicaron que la proporción de Ni: Co correspondiente a 90:10 muestran efectos optimizados de contraste óptico entre los estados transparente y coloreado. Así mismo presentan los más altos valores de transmitancia óptica para su estado transparente. Los resultados que se discutirán en adelante corresponden principalmente a muestras con esta proporción.

En la fig. 5 se muestra el patrón de difracción de rayos X para una muestra correspondiente a la proporción Ni: Co de 90:10. De ella se puede deducir que en la película están presentes las fases NiO y Co_3O_4 , aunque no podemos dar más detalles microestructurales de los mismos. Un efecto notable de la introducción de nitrato de cobalto en la solución acuosa de nitrato de níquel es la reducción substancial de la rugosidad superficial del recubrimiento.

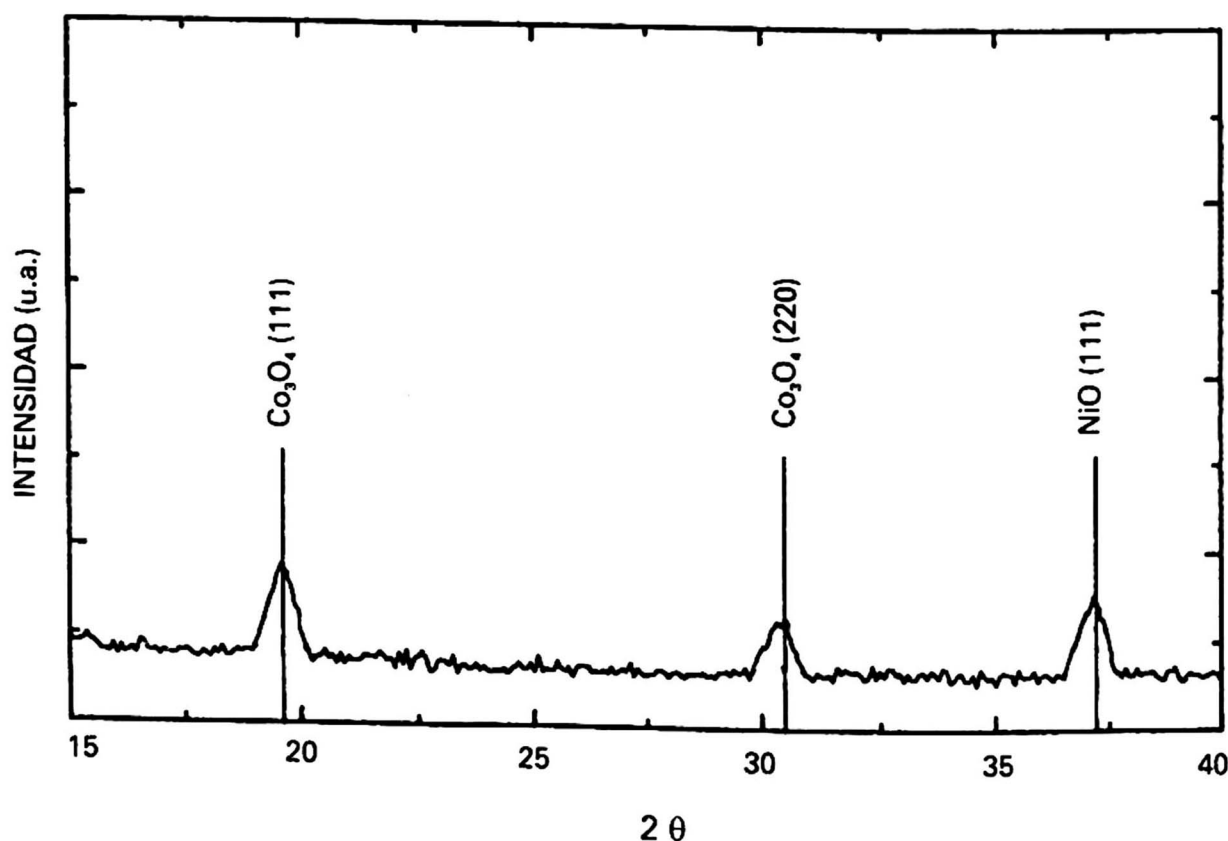


Fig. 5

Patrón de difracción de rayos X de una película delgada de óxido de níquel y cobalto en la proporción 90:10.

La fig. 6 muestra la micrografía óptica para dos superficies una (fig. 6a) correspondiente al óxido de níquel obtenido usando una solución acuosa de nitrato de níquel, y otra (fig. 6b) usando una solución acuosa conteniendo nitrato de níquel

y nitrato de cobalto en la proporción 90:10 respectivamente. De ella podemos notar que detalles morfológicos irregulares que son visiblemente notorios en la fig. 6a, prácticamente no se observan en la fig. 6b. La adición de nitrato de cobalto en la solución permite que esta "moje" más eficientemente el substrato permitiendo un mejor cubrimiento sobre este. En efecto, experimentos realizados con soluciones acuosas de nitrato de cobalto permitieron obtener películas de óxido de cobalto cuyas superficies eran muy uniformes y regulares, aunque su electrocromismo fue muy pobre (menos de 10% de contraste óptico). Por consiguiente se presume que en la mezcla acuosa de nitrato de níquel y nitrato de cobalto se producen dos factores que compiten: el buen electrocromismo y la alta rugosidad superficial del NiO_x con el pobre electrocromismo y la buena regularidad superficial del CoO_x . Por consiguiente la proporción óptima de Ni: Co, que combina apropiadamente los factores que compiten, estaría alrededor de 90:10.

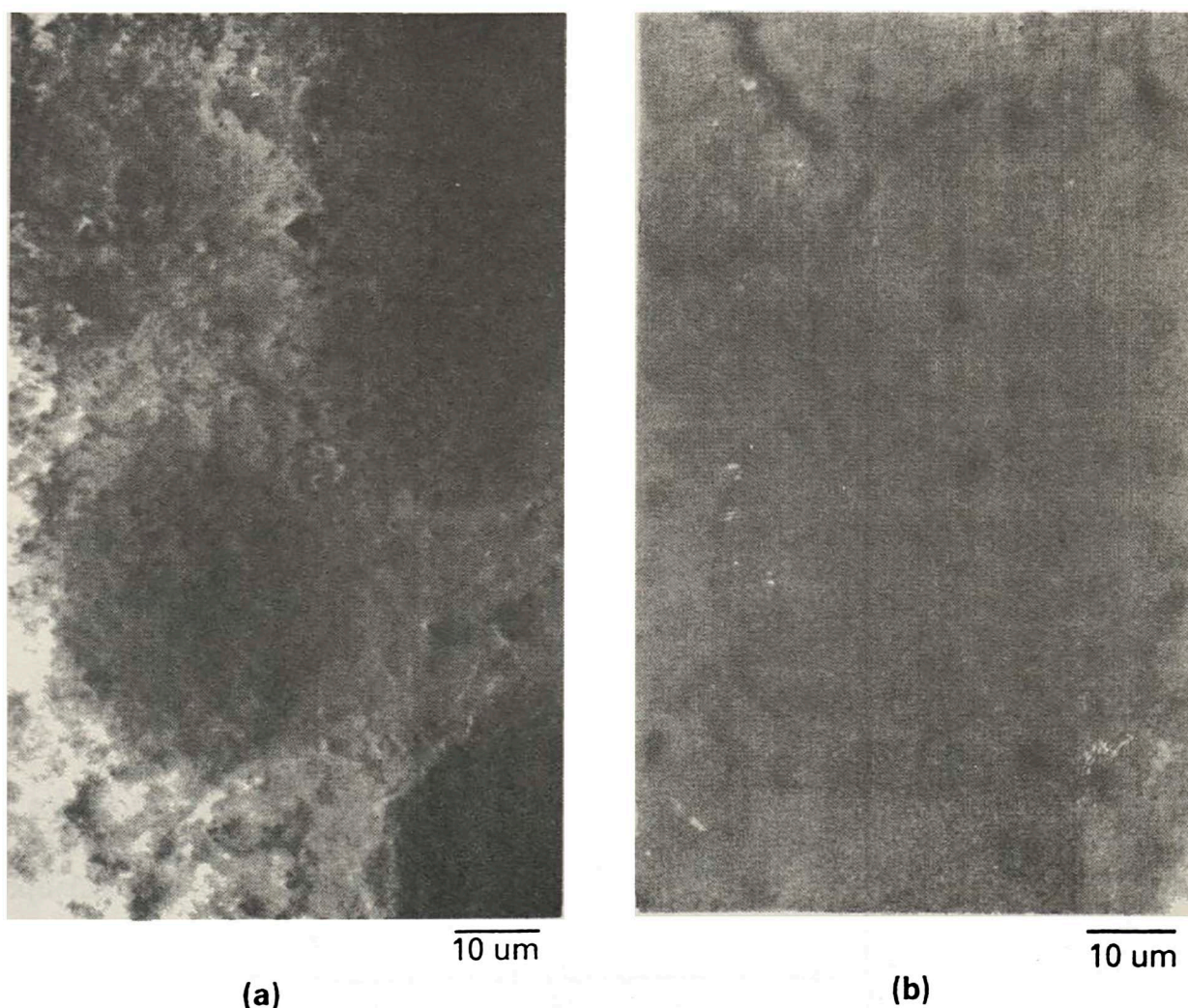


Fig. 6

Micrografías ópticas de películas delgadas de óxido de níquel preparadas en diferentes condiciones. (a) Corresponde al óxido de níquel obtenido usando una solución acuosa de nitrato de níquel, y (b) usando una solución acuosa conteniendo nitrato de níquel y nitrato de cobalto en la proporción 90:10 respectivamente.

La fig. 7 muestra los resultados de las medidas in-situ de la transmitancia óptica (parte superior) y la voltametría cíclica (parte inferior), para el estado estable de inserción/extracción reversible de carga. Podemos observar un electrocromismo pronunciado (un contraste óptico de más de 50%), apreciándose el carácter anódico del proceso de coloración. El perfil voltamétrico difiere ligeramente al mostrado en la fig. 2 correspondiente al NiO_x , particularmente en la zona catódica. Esto probablemente se deba a la influencia electroquímica del óxido de cobalto, presente en la película.

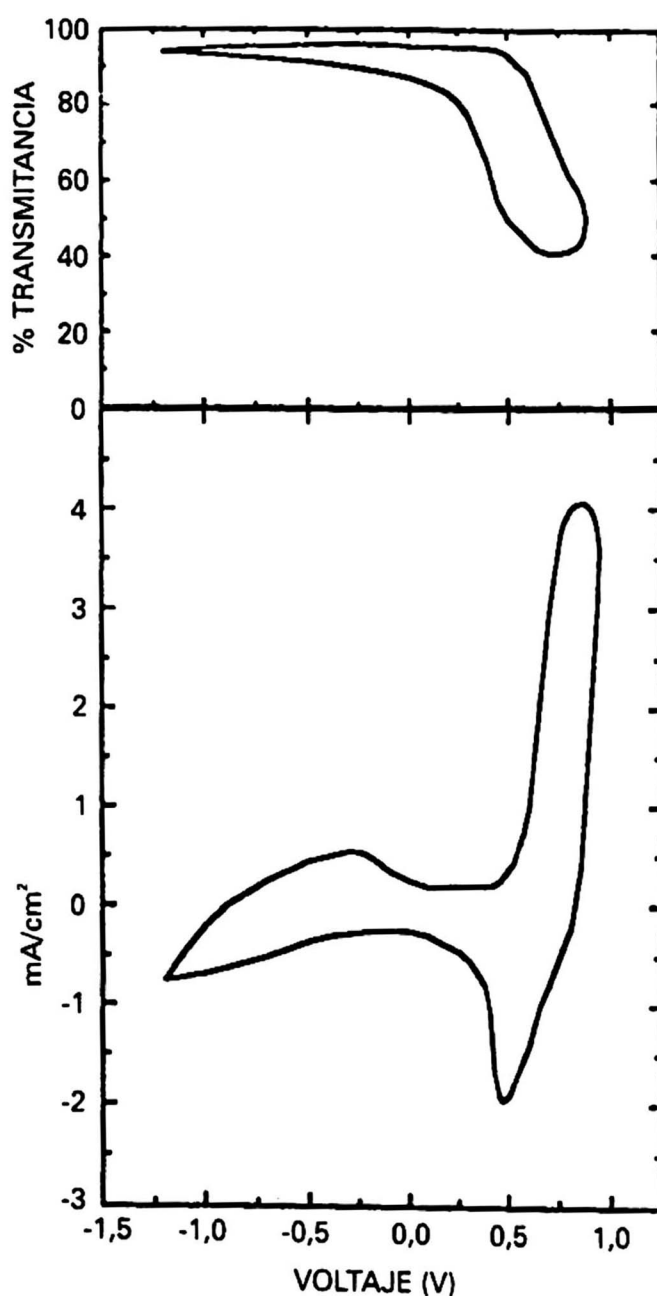


Fig. 7

Voltametría cíclica (parte superior) asociada a la transmitancia in-situ (parte inferior) de una película delgada de óxido de níquel y cobalto (Ni: Co) para una proporción de 90: 10 respectivamente. El sustrato es vidrio precubierto de ITO y el tiempo de depósito de 30 min.

Por otro lado debemos resaltar la alta transmitancia para el estado transparente, casi 95% para $\lambda = 630$ nm, el mismo que se corrobora en la fig. 8 en donde se muestra los valores de la transmitancia espectral para los estados de mayor transparencia y de mayor coloración, respectivamente.

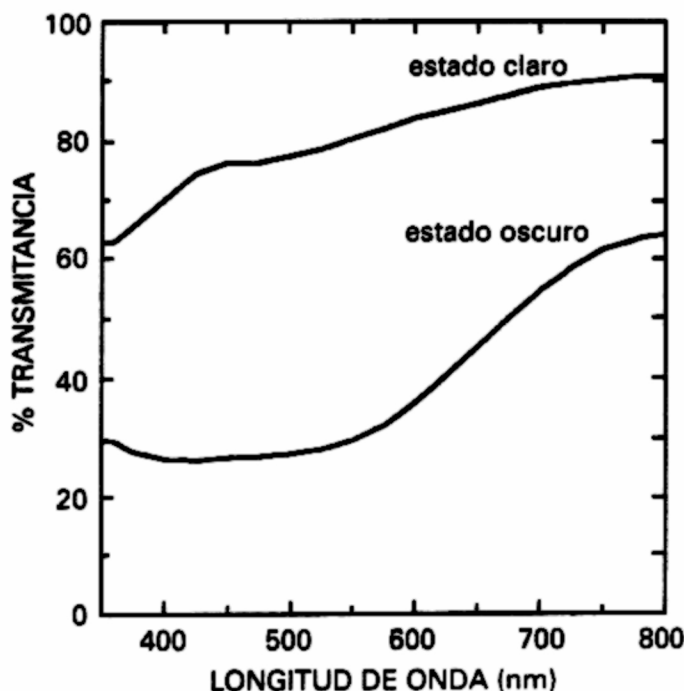


Fig. 8

Transmitancia espectral óptica para los estados claro y oscuro de una película delgada de óxido de níquel y cobalto, preparada de una solución acuosa de $\text{Ni}(\text{NO}_3)_2 \cdot 6\text{H}_2\text{O}$ y $\text{Co}(\text{NO}_3)_3 \cdot 6\text{H}_2\text{O}$ en una proporción molar de 90:10 respectivamente. Substratos de vidrio precubierto con ITO y con un tiempo de depósito de 30 min.

Esta notable mejora del comportamiento óptico, respecto a aquellas obtenidas de soluciones acuosas de nitrato de níquel [Arakaki et. al.], está probablemente ligada al dramático cambio de la calidad superficial del recubrimiento con la adición de pequeñas porciones de nitrato de cobalto en la solución pirolítica.

Conclusión

Puede conseguirse un buen comportamiento electrocrómico en las películas de óxido de níquel pirolítico utilizando cualquiera de los procedimientos discutidos: empleando soluciones alcohólicas de nitrato de níquel o añadiendo a la solución acuosa de nitrato de níquel cantidades ~10% de nitrato de cobalto.

Este resultado indica que usando la técnica del rocío pirolítico o "spray pyrolysis", técnica sencilla y de bajo costo, es posible obtener óxido de níquel electrocrómico con propiedades similares a aquellas obtenidas con técnicas más sofisticadas y de mayor costo, como la técnica de "sputtering".

REFERENCIAS

1. Arakaki, J.; Reyes, R.; Horn, M. and Estrada, W. *Solar Energy Matter.*, 37 (1995) 33.
2. Carpenter, M.; Conell, and D. A. Corrigan, D. A. *Solar Energy Matter*, 16 (1987) 333.
3. Carpenter, M. and Corrigan, D. (Eds.), *Proc. Symp. Electrochromic Materials, The Electrochem. Soc.* 90 (2) 1990.
4. Fantini, M. and Gorenstein, A. *Solar Energy Matter*, 16 (1987) 487.
5. Granqvist, C. G., *Handbook of Inorganic Electrochromic Materials*, Elsevier Science, 1995.
6. Lampert, C. M. and Granqvist, C. G. (Eds.), *Large Area Chromogenics: Materials and Devices for Transmission Control*, SPIE Press, Vol. IS4, 1988.
7. Svensson, J. S. E. M. and Granqvist, C. G. *Solar Energy Matter.*, 16 (1987) 19.

FE DE ERRATA (Volumen 1, número 1, junio 1995)

El problema de la integral de Feynman

H. G. Valqui (*)

ABSTRACT

The Schroedinger equation controls the physical system dynamics, by means of a probability amplitude function (which expresses the probability of the numerical values of the physical observables). Such function is well determined at time t if one knows it at initial time t_0 . Feynman built up a probabilistic operator -called the propagator- which applied to the function at time t_0 gives directly the corresponding function at time t . To that end, Feynman had to realize a certain "summation" including all possible differentiable trajectories.

In this paper, after a short introduction, it is shown how to achieve that "summation", and it is calculated in the particular case of the free particle. After that, the Schroedinger equation is derived, as well as a formula for the approximated calculation of the propagator, when a potential function is present.

RESUMEN

La ecuación de Schroedinger determina la dinámica de los sistemas físicos, por medio de una función de amplitud de probabilidad (la probabilidad de que los observables físicos tomen valores específicos). Dicha función quedará bien determinada en todo instante t , si ella es conocida en un instante inicial t_0 .

(*) Facultad de Ciencias, Universidad Nacional de Ingeniería, Casilla postal 31-139, Lima Perú
e-mail: valqui@fc-uni.pe.

Feynman construyó un operador probabilístico, llamado propagador, que aplicado a la función inicial reproduce directamente la función en el instante t . Para ello, Feynman propuso realizar una cierta "suma" sobre todas las trayectorias posibles, continuamente diferenciables.

En el presente trabajo, después de una introducción, se muestra la construcción de dicha "suma" y se la realiza en el caso particular de la partícula libre. Luego se obtiene la ecuación de Schroedinger y se construye una fórmula para el cálculo aproximado del propagador para el caso en el que el potencial es no nulo.

Introducción

Para describir la dinámica de los sistemas físicos sometidos a un potencial $V(t, x)$, la física clásica postula la existencia de una función lagrangeana

$$L: \mathbb{R}^{1+2n} \rightarrow \mathbb{R}$$

con la cual se construye la funcional de Acción S ,

$$S(t_2, a_2, t_1, a_1; q) = \int_{t_1}^{t_2} dt L(t, q(t), q'(t))$$

correspondiente a cada trayectoria $q: \mathbb{R}^n \rightarrow \mathbb{R}$, que conecta a los "puntos" cronoespaciales (t_1, a_1) y (t_2, a_2) , previamente fijados, donde $t_1 \leq t_2$.

De entre todas las trayectorias posibles, que conecten los puntos indicados, sólo aquella (o aquellas) que proporcionen a S un valor extremal resultará físicamente realizada (Principio de Hamilton). Tal condición extremal implica que la función L deba satisfacer la ecuación diferencial de Euler-Lagrange (cuyas soluciones son precisamente las trayectorias físicamente realizables).

En cambio, en la física cuántica, dado el potencial $V(t, x)$, se postula la existencia de:

- 1> una función compleja,

$$\Phi: \mathbb{R}^{1+n} \rightarrow \mathbb{C}$$

que debe contener toda la información sobre el sistema físico en consideración;

- 2> un operador diferencial espacial, llamado hamiltoniano, que en el caso no relativístico tiene la forma

$$H = -(\hbar^2/2m) \nabla^2 + V \quad \{3\}$$

donde ∇^2 es el laplaceano n-dimensional y V es operador de potencial actuando como un operador multiplicativo (operador diagonal, $V\Psi = v\Psi$); \hbar y m son ciertas constantes;

- 3> una ecuación diferencial, llamada ecuación de Schroedinger,

$$H \Phi + (\hbar/i) \partial \Phi = 0 \quad (02)$$

donde ∂ es el operador de derivada temporal. Las soluciones contienen, en forma probabilística, con densidad de probabilidad igual a $|\Phi|^2$, toda la información sobre el sistema físico considerado.

El propagador de Feynman

En la ecuación de Schroedinger, se debe dar en el instante t_0 , la función inicial $\Phi(t_0)$ para obtener, en el instante t, la función $\Phi(t)$. Por ello, Feynman se propuso construir un operador llamado -propagador temporal- $K(t, t_0)$ con $t_0 \leq t$, que al ser aplicado sobre el vector $\Phi(t_0)$ lo transforme en el vector $\Phi(t)$,

$$K(t, t_0) \Phi(t_0) = \Phi(t) \quad (03)$$

o, en forma desarrollada,

$$\int_{-\infty}^{\infty} dz K(t, x, t_0, z) \Phi(t_0, z) = \Phi(t, x)$$

(donde la integración se realiza en todo el espacio considerado. Debe notarse que la expresión anterior representa el producto de una matriz de índices continuos por un vector de índice también continuo).

El mencionado propagador debería incluir el efecto de todas las posibles trayectorias continuamente diferenciables (y no sólo de la trayectorial extremal). Con tal fin, Feynman postuló

$$K(t, x, t_0, x_0) = \sum_q e^{(i/\hbar) S(t, t_0, x, x_0; q)} \quad (04)$$

donde la "suma" debería realizarse -de alguna manera- según todas las trayectorias continuamente diferenciables.

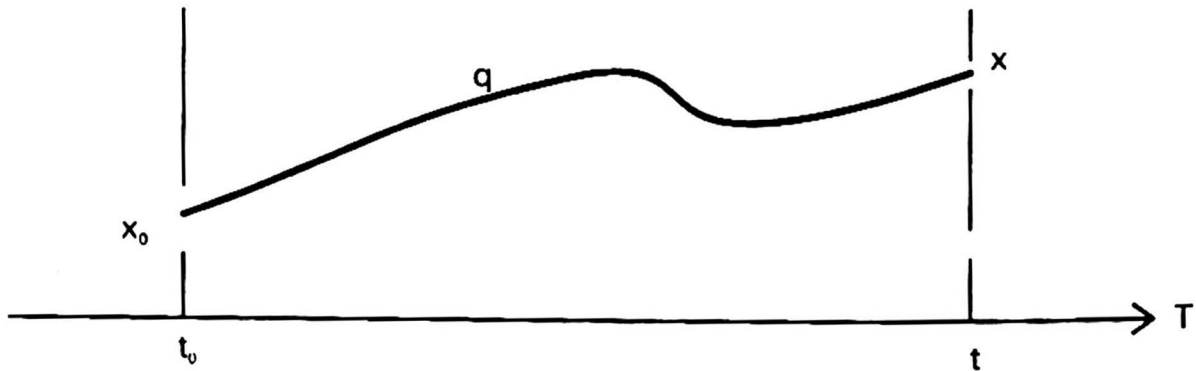


Fig. 1 Una trayectoria q , que en el instante t_0 parte del punto x_0 , para llegar, en el instante t , al punto x .

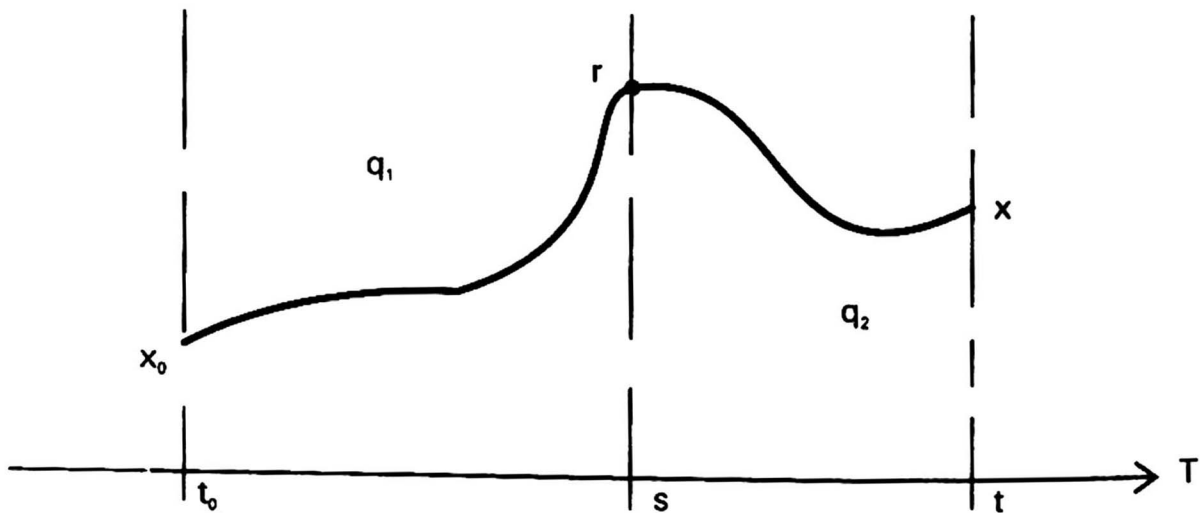


Fig. 2. Una trayectoria q^r , que conecta a los puntos crono-espaciales inicial y final, pero que en el instante dado s debe pasar por el punto espacial r .

Fijemos un instante intermedio s , $t_0 < s < t$ y un punto espacial r . Consideremos todas las trayectorias q^r , que conectando los puntos inicial y final pasen, en el instante s , por el punto espacial r , como se muestra en la fig. 2.

Sea q_1 una trayectoria cualquiera que conecta a los puntos crono-espaciales (t_0, x_0) y (s, r) ; sea q_2 una trayectoria que conecta a los puntos (s, r) y (t, x) . Entonces, cualquier q^r puede ser considerada como la reunión de una cierta trayectoria q_1 y una cierta trayectoria q_2 . Ahora, cualquier trayectoria q puede ser obtenida eligiendo, primeramente y para un instante s , un cierto punto espacial r , luego, eligiendo las correspondientes trayectorias q_1 y q_2 . Para una trayectoria $q^r = q_1 \cup q_2$ podemos escribir,

$$S(t, b, t_0, a; q^r) = S(t, b, s, r; q_2) + S(s, r, t_0, a; q_1)$$

y, también

$$e^{\tau \cdot S(t, b, t_0, a; q)} = e^{\tau \cdot S(t, b, s, r; q_2)} \cdot e^{\tau \cdot S(s, r, t_0, a; q_1)}$$

Es decir, en general, para incluir todas las trayectorias, podemos escribir,

$$\sum_q F(t, b, t_0, a; q) = \int_{-\infty}^{\infty} dr \left[\sum_{q_2} \sum_{q_1} F(t, b, s, r; q_2) \cdot F(s, r, t_0, a; q_1) \right] \quad (05)$$

Aplicando (05) a nuestro caso obtendremos,

$$K(t, x, t_0, x_0) = \int_{-\infty}^{\infty} dr K(t, x, s, r) K(s, r, t_0, x_0)$$

o, en forma abreviada

$$K(t, t_0) = K(t, s) K(s, t_0) \quad (06)$$

donde s es cualquier instante intermedio.

Considerando la propiedad (06) dividamos el intervalo (t_0, t) en N subintervalos, con $t_N = t$; entonces podemos escribir,

$$K(t, t_0) = \prod_{j=0}^{N-1} K(t_{j+1}, t_j) \quad (07)$$

o, también,

$$K(t, x, t_0, x_0) = \int dx_{N-1} dx_{N-2} \dots dx_j \dots dx_1 \prod_{i=0}^{N-1} K(t_{i+1}, x_{i+1}, t_i, x_i)$$

donde

$$K(t_{j+1}, x_{j+1}, t_j, x_j) = \sum_{q_j} e^{(i/\hbar) S(t_{j+1}, x_{j+1}, t_j, x_j; q_j)} \quad (08)$$

$$S(t_{j+1}, x_{j+1}, t_j, x_j; q_j) = \int_{t_j}^{t_{j+1}} dt L(t, q_j(t), q'_j(t)) \quad (09)$$

Notemos que, en principio, algunas trayectorias q_j pueden tomar, para instantes t intermedios entre t_j y t_{j+1} , valores $q_j(t)$ muy alejados tanto de x_j como de x_{j+1} . Pero, por otra parte, teniendo presente la continuidad de las trayectorias, cuando la diferencia $\varepsilon_j = t_{j+1} - t_j$ sea muy pequeña, los valores x_j , x_{j+1} y $q_j(t)$ deberán ser casi coincidentes. Es decir, en tales circunstancias, todas las trayectorias q_j serán casi coincidentes. Por la continuidad de la derivada, también las velocidades q'_j serán casi coincidentes.

Ahora, para los casos en los que la función lagrangeana sea continua, y para valores de ε_j suficientemente pequeños, podemos aplicar en (09) el teorema del valor intermedio,

$$S(t_{j+1}, x_{j+1}, t_j, x_j; q_j) = \varepsilon_j L(\theta, q_j(\theta), q'_j(\theta)), \quad t_j \leq \theta \leq t_{j+1}$$

Considerando lo argumentado anteriormente, sobre la casi coincidencia de las trayectorias y las velocidades, una buena aproximación para los valores intermedios sería,

$$\theta \approx t_j + \varepsilon_j/2, \quad q_j(\theta) \approx (x_{j+1} + x_j)/2, \quad q'_j(\theta) \approx (x_{j+1} - x_j)/\varepsilon_j$$

[con lo cual a todas las trayectorias entre (t_j, x_j) y (t_{j+1}, x_{j+1}) les corresponde un único valor intermedio; lo mismo sucede con las velocidades].

Eligiendo, por comodidad, todos los $\varepsilon_j = \varepsilon = (t - t_0)/N$, podremos ensayar la aproximación (postulado de Feynman)

$$K(t_{j+1}, x_{j+1}, t_j, x_j) = \sum_{q_j} e^{(i/h) S_j} \approx A(\varepsilon) e^{(i\varepsilon/h) L(t_j + \varepsilon/2, (x_{j+1} + x_j)/2, (x_{j+1} - x_j)/\varepsilon)} \quad (10)$$

donde $A(\varepsilon)$ es un coeficiente que, por una parte, expresaría la multiplicidad de las trayectorias q_j y por otra parte, debe permitir la existencia del límite cuando ε tienda a cero,

$$K(t, x, t_0, x_0) = \lim_{\varepsilon \rightarrow 0} A(\varepsilon)^N \int dx_{N-1} \dots dx_1 \prod_{j=0}^{N-1} e^{(i\varepsilon/h) L_j} \quad (11)$$

donde, $L_j \equiv L(t_j + \varepsilon/2, (x_{j+1} + x_j)/2, (x_{j+1} - x_j)/\varepsilon)$.

Esta es una forma general para el propagador de Feynman, donde el coeficiente $A(\varepsilon)$ todavía es una incógnita.

Un caso particular: La partícula libre

Consideremos el caso de la partícula libre, donde $V = 0$. Entonces,

$$L(t_j + \varepsilon/2, (x_{j+1} + x_j)/2, (x_{j+1} - x_j)/\varepsilon) = (m/2) ((x_{j+1} - x_j)/\varepsilon)^2$$

Para $\alpha = m/(2\hbar\epsilon)$,

$$\Gamma(\epsilon) = \int dx_{N-1} \dots dx_1 e^{i\alpha \sum_{j=0}^{N-1} (x_{j+1} - x_j)^2}$$

el propagador quedará expresado por

$$K_0(t, x, t_0, x_0) = \lim_{\epsilon \rightarrow 0} A(\epsilon)^N \Gamma(\epsilon) \quad (12)$$

Ahora, teniendo presente la identidad algebraica

$$(1/\mu)(y-x)^2 + (x-z)^2 = (1 + 1/\mu)(x-x^0)^2 + (1/(\mu+1))(y-z)^2,$$

con $x^0 = (y + \mu z)/(\mu + 1)$, podemos escribir,

$$\begin{aligned} \sum_{j=0}^{N-1} (x_{j+1} - x_j)^2 &= (x_1 - x_1^0)^2 + (1/2)(x_2 - x_0)^2 + \sum_{j=2}^{N-1} (x_{j+1} - x_j)^2 \\ &= \sum_{j=1}^{N-1} (1 + 1/j)(x_j - x_0^0)^2 + (1/N)(x - x_0)^2 \end{aligned}$$

de donde obtenemos

$$\Gamma(\epsilon) = e^{i(\alpha/N)(x-x_0)^2} \int dx_{N-1} \dots dx_1 \prod_{j=1}^{N-1} e^{i\alpha(1+1/j)(x_j - x_0^0)^2}$$

Pero, en el caso unidimensional, las integrales en el plano complejo (ver apéndice) tienen la misma expresión que en la recta real,

$$\int_{-\infty}^{\infty} dx e^{iax^2} = \sqrt{(i\pi/a)} \quad (13)$$

entonces,

$$\prod_{j=1}^{N-1} \int dx_j e^{i\alpha(1+1/j)(x_j - x_0^0)^2} = \prod_{j=1}^{N-1} [i\pi \cdot j/\alpha(j+1)]^{1/2}$$

de donde,

$$\Gamma(\epsilon) = e^{i(\alpha/N)(x-x_0)^2} \sqrt{1/N} (i\pi/\alpha)^{(N-1)/2}$$

lo que nos permite escribir,

$$A(\epsilon)^N \Gamma(\epsilon) = \sqrt{\alpha/(i\pi N)} A(\epsilon)^N (i\pi/\alpha)^{N/2}$$

o, teniendo presente que $N \cdot \varepsilon = t - t_0$,

$$A(\varepsilon)^N \Gamma(\varepsilon) = \sqrt{m/(i 2\pi \hbar (t - t_0))} A(\varepsilon)^N (i 2\pi \hbar \varepsilon/m)^{N/2}$$

Entonces, con el objeto de que la expresión anterior tenga límite cuando ε tiende a cero, resulta conveniente asignar al coeficiente $A(\varepsilon)$ el valor

$$A(\varepsilon) = \sqrt{m/(i 2\pi \hbar \varepsilon)} \quad (14)$$

con lo cual, finalmente obtenemos

$$K_0(t, x, t_0, x_0) = \sqrt{m/(i 2\pi \hbar (t - t_0))} e^{i(m/2\hbar)(x - x_0)^2/(t - t_0)} \quad (15)$$

que es la expresión para el propagador de la partícula libre.

El propagador para un potencial independiente del tiempo

Definiendo el operador hamiltoniano

$$H(t) = (-\hbar^2/2m) \nabla^2 + V(t),$$

donde $V(t)$ es un operador multiplicativo (diagonal)

$$V(t) \Phi(t) = V(t) \Phi(t)$$

la ecuación de Schroedinger puede ser escrita

$$H(t) \Phi(t) + (\hbar/i) \partial \Phi(t) = 0 \quad (16)$$

(donde el tiempo puede ser creciente o decreciente; si se desea que la ecuación sea válida sólo para tiempos crecientes deberá introducirse tal condición en forma explícita).

Para el caso en el que el potencial V no dependa del tiempo, escribiremos

$$H_0 = (-\hbar^2/2m) \nabla^2 + V;$$

En este caso particular, se cumplirá que

$$\partial \Phi(t) = (-i/\hbar) H_0 \Phi \Rightarrow \partial^n \Phi(t) = (-i/\hbar)^n (H_0)^n \Phi(t)$$

es decir,

$$e^{\tau \partial} \Phi(t) = e^{(-i/\hbar) \tau \cdot H_0} \Phi(t) \quad (17)$$

Por otra parte, suponiendo la analiticidad temporal de la función Φ al desarrollarla en su serie de Taylor obtendremos

$$\Phi(t + \tau) = [e^{\tau \cdot \partial}] \cdot \Phi(t),$$

o también,

$$\Phi(t + \tau) = U(\tau) \cdot \Phi(t),$$

donde, el operador

$$U(\tau) = e^{(-i/\hbar)\tau \cdot H_0} \quad (18)$$

sólo contiene operadores espaciales. Este operador $U(\tau)$, debido a su efecto sobre la función Φ , es llamado operador de propagación temporal (de las funciones que pertenecen al espacio de soluciones de la ecuación de Schroedinger).

Comparando este resultado con el del propagador K_0 , para el caso en el que V no dependa del tiempo, podemos deducir que ambos operadores son coincidentes,

$$\partial V = 0 \Rightarrow U(\tau) = K_0(t + \tau, t) \quad (19)$$

(eventualmente, para asegurar que K_0 actúe sólo para tiempos crecientes, el operador U debe ser multiplicado por la función de escalón temporal).

Claro está que, entonces, los elementos de la matriz continua U serán

$$U(\tau, x, y) = K_0(t + \tau, x, t, y) \quad (20)$$

o, para t crecientes

$$\theta(\tau) \cdot U(\tau, x, y) = K_0(t + \tau, x, t, y) \quad (20a)$$

Por otra parte, ya hemos visto que para τ suficientemente pequeño, por el postulado de Feynman, (10), podemos escribir

$$U(\tau, x, y) = K_0(t + \tau, x, t, y) \approx A(\tau) e^{-\pi \cdot A(\tau)^2 \cdot (y-x)^2 - (i\tau/\hbar) V((x+y)/2)}$$

y también,

$$\Phi(t + \tau, x) = [K_0(t + \tau, t) \cdot \Phi(t)](x) \approx A(\tau) \cdot \int dy \cdot e^{-\pi \cdot A(\tau)^2 \cdot (y-x)^2 - (i\tau/\hbar) V} \Phi(t, y)$$

donde, en el límite, \approx se convierte en un signo de igualdad. Cambiando la variable de integración $y = x + z$, con lo cual no se alteran los límites de integración, [por simplicidad consideraremos $V = 0$].

$$\Phi(t + \tau, x) \approx A(\tau) \cdot \int dz \cdot e^{-\pi \cdot A(\tau)^2 \cdot z^2} \Phi(t, x + z)$$

Al desarrollar $\Phi(t, x + z)$ en serie de Taylor, las correspondientes potencias impares se anularán en la integración, y las potencias pares, al ser integradas producirán sumandos con coeficientes proporcionales a potencias positivas de τ (ver apéndice). Sólo el primer sumando, después de multiplicar por $A(\tau)$, resulta independiente de τ ; de manera que al pasar al límite, cuando τ tiende a cero, se anularán todos los sumandos, excepto el primero. Es decir, obtenemos

$$\Phi(t + \tau, x) = [K_0(t + \tau, t) \cdot \Phi(t)](x) \xrightarrow{\tau \rightarrow 0} \Phi(t, x)$$

con lo cual se verifica que

$$[K_0(t + \tau, t)]_{x,y} = [e^{(-i/\hbar)\tau \cdot H_0}]_{x,y} \xrightarrow{\tau \rightarrow 0} \delta(x - y)$$

o, si se prefiere,

$$K_0(t + \tau, t) = e^{(-i/\hbar)\tau \cdot H_0} \xrightarrow{\tau \rightarrow 0} I$$

es decir, en el límite, como debería ser, $K_0(t, \tau)$ se reduce a la matriz unidad.

La ecuación de Schroedinger

Calculemos la derivada temporal de la función $\Phi(t)$. Para ello consideremos la diferencia

$$\Phi(t + \varepsilon) - \Phi(t) = K(t + \varepsilon, t) \Phi(t) - \Phi(t),$$

o, en forma desarrollada,

$$\Phi(t + \varepsilon, x) = \int dy K(t + \varepsilon, x, t, y) \Phi(t, y)$$

Usando la aproximación (10), podemos escribir

$$\Phi(t + \varepsilon, x) \approx A(\varepsilon) \int dy e^{(i\varepsilon/\hbar)L(t + \varepsilon/2, (x+y)/2, (y-x)/2)} \Phi(t, y)$$

Recurriendo al cambio de variable $y = z + x$, con lo cual no resultan afectados los límites de integración,

$$\Phi(t + \varepsilon, x) \approx A(\varepsilon) \int dz e^{(i\varepsilon/\hbar)L(t + \varepsilon/2, x + z/2, z/\varepsilon)} \Phi(t, x + z)$$

Para el lagrangeano $L(t, x, w) = (m/2) w^2 - V(t, x)$, con el α definido anteriormente, podemos escribir,

$$\Phi(t + \varepsilon, x) \approx A(\varepsilon) \int dz e^{i\alpha z^2 - i(\varepsilon/h)V(t + \varepsilon/2, x + z/2)} \Phi(t, x + z) \quad (21)$$

Con el objeto de calcular la derivada temporal, basta desarrollar las funciones del segundo miembro de (21) en la primera potencia de ε . Notemos que no es necesario desarrollar el potencial, pues él ya está afectado de un coeficiente ε . Por otra parte, es claro que z tenderá a cero cuando ε tienda a cero, pero todavía ignoramos la forma de tal variación.

Ahora, desarrollemos Φ en serie de Taylor. Entonces, considerando, por simplicidad, el caso unidimensional,

$$\Phi(t, x + z) = \Phi(t, x) + z \Phi_1(t, x) + (z^2/2) \Phi_2(t, x) + \dots + z^n \mu(t, x, z)$$

Además, teniendo presente que ε es muy pequeño, podemos escribir, dentro del grado de aproximación indicado,

$$e^{-i(\varepsilon/h)V(t + \varepsilon/2, x + z/2)} \Phi(t, x + z) = [1 - i(\varepsilon/h)V(t, x)] [\Phi + z \Phi_1 + (z^2/2) \Phi_2 + \dots]$$

donde, a la derecha aparecen $2n + 2$ sumandos. Sólo el primero de dichos sumandos es independiente tanto de ε como de z .

Ahora reescribiremos la expresión (21),

$$\Phi(t + \varepsilon, x) \approx A(\varepsilon) \int dz e^{i\alpha z^2} \{ \Phi(t, x) + 2n + 1 \text{ sumandos } (\varepsilon, z) \}$$

o, realizando la integración,

$$\Phi(t + \varepsilon, x) \approx A(\varepsilon) \sqrt{2\pi i h \varepsilon / m} \Phi(t, x) + A(\varepsilon) \int dz e^{i\alpha z^2} \{ 2n + 1 \text{ sum. } (\varepsilon, z) \}$$

Aquí nuevamente podemos apreciar la conveniencia de asignar al coeficiente $A(\varepsilon)$ el valor dado en (14). Entonces obtenemos la diferencia necesaria para la derivación temporal,

$$\Phi(t + \varepsilon, x) - \Phi(t, x) \approx A(\varepsilon) \int dz e^{i\alpha z^2} \{ 2n + 1 \text{ sumandos } (\varepsilon, z) \}$$

De los $2n + 1$ sumandos aquí considerados, los que son proporcionales a las potencias impares de z se anularán al realizar la integración.

Por otro lado, para las potencias pares tenemos que (ver apéndice),

$$\int_{-\infty}^{\infty} dx x^{2k+2} e^{i\alpha x^2} = i(2k+1)/(2\alpha) \int_{-\infty}^{\infty} dx x^{2k} e^{i\alpha x^2} \quad (22)$$

para $k = 0, 1, 2, \dots$, lo cual, teniendo presente el valor dado en (13), nos permite escribir

$$\int_{-\infty}^{\infty} dx x^{2k} e^{i\alpha x^2} = (2k-1)!! \cdot (i/2\alpha)^k \cdot \sqrt{i\pi/\alpha} \sim \alpha^{-k-1/2} \sim \varepsilon^{k+1/2}$$

donde $(2k-1)!! = 1 \cdot 3 \cdot 5 \cdot \dots \cdot (2k-1)$. Después de multiplicar por $A(\varepsilon)$ la diferencia resulta proporcional a una serie de potencias de ε con exponentes $k = 1, 2, \dots$. Entonces, al dividir entre ε y pasar al límite $\varepsilon \rightarrow 0$, sólo el primer sumando de la derecha no se anulará, quedando

$$\partial \Phi(t, x) = (-i/\hbar) V(t, x) \Phi(t, x) - \hbar/(2im) \Phi_2$$

o, también

$$(-\hbar^2/2m) \nabla^2 \Phi(t, x) + V(t, x) \Phi(t, x) + (\hbar/i) \partial \Phi(t, x) = 0$$

o,

$$(-\hbar^2/2m) \nabla^2 \Phi(t) + V(t) \Phi(t) + (\hbar/i) \partial \Phi(t) = 0 \quad (23)$$

es decir, el vector $\Phi(t)$, obtenido aplicando el propagador al vector arbitrario $\Phi(t_0)$, resulta ser solución de la ecuación de Schroedinger, correspondiente a dicho valor inicial $\Phi(t_0)$.

Finalmente, teniendo presente que $\Phi(t) = K(t, t_0) \Phi(t_0)$, podemos escribir,

$$(-\hbar^2/2m) \nabla^2 K(t, t_0) \Phi(t_0) + V(t) K(t, t_0) \Phi(t_0) + (\hbar/i) \partial K(t, t_0) \Phi(t_0) = 0$$

donde $\Phi(t_0)$ es un vector arbitrario; entonces, obtenemos

$$(-\hbar^2/2m) \nabla^2 K(t, t_0) + V(t) K(t, t_0) + (\hbar/i) \partial K(t, t_0) = 0 \quad (24)$$

es decir, este propagador de Feynman también es solución de la ecuación de Schroedinger.

En la deducción anterior no se ha considerado que el operador $K(t, t_0)$ actúe sólo para tiempos crecientes; ésta es una exigencia que es necesario imponer para garantizar su carácter probabilístico. En tal caso, considerando el caso particular en

el que V no depende del tiempo, y teniendo presente (20a), la ecuación de Schroedinger toma la forma

$$(-\hbar^2/2m) \nabla^2 K_0(t, t_0) + V \cdot K_0(t, t_0) + (\hbar/i) \partial K_0(t, t_0) = i \cdot \hbar \cdot \delta(t - t_0)$$

lo que nos permite afirmar que, también en el caso general, la ecuación que asegura el efecto "hacia el futuro" del propagador es de la forma

$$(-\hbar^2/2m) \nabla^2 K(t, t_0) + V(t) \cdot K(t, t_0) + (\hbar/i) \partial K(t, t_0) = i \cdot \hbar \cdot \delta(t - t_0)$$

o, si se prefiere

$$H(t) \cdot K(t, t_0) + (\hbar/i) \partial K(t, t_0) = i \cdot \hbar \cdot \delta(t - t_0) \quad (24a)$$

El caso de perturbación dependiente del tiempo

Consideremos el caso en que el hamiltoniano $H(t)$ puede expresarse como la suma

$$H(t) = H_0 + W(t), \text{ con } H_0 = (-\hbar^2/2m) \nabla^2 + V, \quad \partial V = 0$$

entonces tendremos las ecuaciones

$$[H_0 + (\hbar/i) \partial] K(t, t_0) + W(t) \cdot K(t, t_0) = i \cdot \hbar \cdot \delta(t - t_0)$$

y, por otra parte,

$$[H_0 + (\hbar/i) \partial] K_0(t, t_0) = i \cdot \hbar \cdot \delta(t - t_0) \quad (25)$$

donde $K_0(t, t_0)$ está expresado por H_0 como se ve de (18) y (20a), y lo suponemos calculable de alguna manera.

Restando las igualdades anteriores obtenemos

$$H_0 \cdot J(t, t_0) + (\hbar/i) \partial J(t, t_0) = -W(t) \cdot K(t, t_0) \quad (26)$$

donde $J(t, t_0) = K(t, t_0) - K_0(t, t_0)$

Por otra parte, expresando (25) entre los instantes t, t_1 y aplicando ambos miembros al producto $W(t_1) K(t_1, t_0)$, tendremos

$$[H_0 + (\hbar/i) \partial] K_0(t, t_1) \cdot W(t_1) \cdot K(t_1, t_0) = i \cdot \hbar \cdot \delta(t - t_1) W(t_1) K(t_1, t_0)$$

e, integrando en el intervalo (t, t_0) , obtenemos

$$[H_0 + (\hbar/i) \partial] \cdot \int_{t_0}^t K_0(t, t_1) \cdot W(t_1) \cdot K(t_1, t_0) = i \cdot \hbar \cdot W(t) K(t, t_0) \quad (27)$$

Podemos apreciar que multiplicando la ecuación (26) por $i\hbar$, ella coincidirá con la ecuación (27). Es decir, la expresión integral es una solución para $J(t, t_0)$. Así obtenemos

$$K(t, t_0) = K_0(t, t_0) + (i/\hbar) \cdot \int_{t_0}^t dt_1 \cdot K_0(t, t_1) \cdot W(t_1) \cdot K(t_1, t_0) \quad (28)$$

que es una ecuación integral para $K(t, t_0)$ en función del operador "conocido" $K_0(t, t_0)$ y del potencial "perturbador" $W(t)$.

Escribamos la ecuación (28) en forma abreviada

$$K = K_0 + (i/\hbar) \langle K_0 \cdot W \cdot K \rangle \quad (28a)$$

Definamos las soluciones aproximadas

$$K^{(0)} = K_0, \quad K^{(1)} = K_0 + (i/\hbar) \langle K_0 \cdot W \cdot K^{(0)} \rangle \text{ y en general}$$

$$K^{(n+1)} = K_0 + (i/\hbar) \langle K_0 \cdot W \cdot K^{(n)} \rangle, \quad n = 1, 2, \dots \quad (29)$$

donde, la sucesión, si es convergente, deberá converger hacia $K(t, t_0)$. En forma práctica, consideraremos que la convergencia se ha obtenido cuando la diferencia entre dos aproximaciones sucesivas sea "suficientemente pequeña". Es decir, para un valor de n tal que

$$K^{(n+1)} - K^{(n)} \approx 0 \quad (30)$$

Debe tenerse presente que los resultados anteriores también son válidos para el caso en el que W sea independiente del tiempo.

APENDICE

- A: En el intervalo $(-\pi/4, 0)$ se cumple $1 + 4\theta/\pi < \cos 2\theta$, así mismo en el intervalo $(0, \pi/4)$ se cumple $1 - 4\theta/\pi < \cos 2\theta$. Entonces, para $a > 0$, $0 < \alpha \leq \pi/4$, integrando la función $g(z) = \exp(-a \cdot z^2)$ a lo largo de un arco de radio R obtenemos que

$$\left| \int_0^{\pm \alpha} dz g(z) \right| < (\pi/4aR) \cdot g(R) \cdot [\exp(aR^2(4\alpha/\pi)) - 1]$$

que tiende a 0 cuando $R \rightarrow \infty$.

- B. Integrando $g(z)$ a lo largo del triángulo curvilíneo con vértices $0, R, R \cdot \exp(i\alpha)$ ó vértices $0, R, R \cdot \exp(-i\alpha)$; luego de tomar el límite para $R \rightarrow \infty$, obtenemos

$$\begin{aligned} \int_0^\infty dx \, g(x \cdot e^{\pm i\alpha}) &= [1/\exp(\pm i\alpha)] \cdot \int_0^\infty dx \, g(x) \\ &= 1/2 \sqrt{\pi/a} / \exp(\pm i\alpha) \end{aligned}$$

lo que para $\alpha = \frac{\pi}{4}$ da $\int_{-\infty}^\infty dx \, e^{iax^2} = \sqrt{i\pi/a}$

C: De $\int_{-\infty}^\infty dx \, g(x) \cdot x^n = \int_{-\infty}^\infty dx \, g(x) \cdot (-x)^n$

podemos ver que cuando n es impar los dos miembros se diferencian sólo en el signo, luego la integral será nula.

En particular obtenemos

$$e^{-iax^2} \Big|_{-\infty}^\infty = -2ia \int_{-\infty}^\infty dx \, x \cdot e^{-iax^2} = 0$$

D: Para $p(n, x) = x^n e^{-iax^2}$ obtenemos

$$p(n+1, x)' = (n+1) p(n, x) - 2ia \cdot p(n+2, x)$$

de donde, para n par, integrando

$$(n+1)/(2ia) \int_{-\infty}^\infty dx \cdot x^n e^{-iax^2} = \int_{-\infty}^\infty dx \cdot x^{n+2} e^{-iax^2}$$

BIBLIOGRAFÍA {85}

R. Feynman-A. Hibbs, *Quantum Mechanics and Path Integrals* Mc Graw-Hill, 1965

El spin no es necesariamente relativístico

H.G. Valqui (*)

ABSTRACT

A prism decomposes an incident light ray in many outgoing rays of different colours. This idea is applied to the case of an electron-ray directed to a Stern-Gerlach's device, from which emerge two electron-rays of different 'colours'. This decomposition is achieved by postulating that the Schroedinger equation is really a two-component (matricial) equation, in such a way that, when no magnetic field is present, it reproduces trivially two times the (scalar) Schroedinger equation. The passage from the trivial matrix equation to the magnetic-field-dependent one, forces the appearance of the so called spin matrices.

RESUMEN

Un rayo de luz que incide sobre un prisma resulta descompuesto en muchos rayos de diferentes colores. Esta idea es aplicada al caso de un rayo de electrones que incide sobre un aparato de Stern-Gerlach, donde se producen dos rayos salientes. Tal descomposición se obtiene asumiendo que la ecuación de Schroedinger es realmente una ecuación (matricial) de dos componentes, de manera que, en ausencia de un campo magnético, dicha ecuación representa la repetición trivial de la ecuación escalar de Schroedinger. El paso de la ecuación matricial trivial al de la ecuación para el caso en que exista un campo magnético implica la aparición de las matrices de Pauli.

(*) Facultad de Ciencias, Universidad Nacional de Ingeniería, Casilla postal 31-139, Lima - Perú.

Introducción

La construcción del operador de spin realizada por Pauli no reproducía, por un factor 2, los resultados esperados. Posteriormente Dirac logró construir un modelo que sí reproducía correctamente los resultados experimentales; pero tal modelo surgió de la ecuación relativística del electrón, por lo que en adelante se supuso que el spin era necesariamente un efecto relativístico.

En un artículo matemático aparecido en 1981 [1], quedó aclarado que el operador del spin no es necesariamente producto de los postulados relativísticos.

Aquí desarrollamos las características principales de tal construcción para el caso del spin 1/2.

El modelo no relativístico

El modelo clásico de la óptica considera que la luz blanca es una especie de superposición de ondas luminosas de diferentes colores. Esta onda blanca mantiene su estructura mientras no atraviese algún medio material que posea la propiedad de separar a las diferentes ondas componentes. El medio clásico de separación de ondas luminosas está constituido por un prisma: Sobre un prisma incide luz blanca, y de él emergen haces de luz de varios colores, en direcciones diferentes a la del haz incidente.

Por analogía podemos suponer que la función de onda de un conjunto de partículas (átomos) que inciden sobre un aparato de Stern-Gerlach es una "superposición" de dos funciones de onda (cada una correspondiente a uno de los "colores" de spin). Esta "superposición" podría ser detectada recién en presencia de un campo magnético no homogéneo (que desempeñaría el papel de un prisma).

Con el objeto de obtener una ecuación de Schroedinger que proporcione dos soluciones diferentes (cada una correspondiente a un "color" de spin), como un primer paso escribiremos doblemente la misma ecuación estacionaria de Schroedinger:

$$H f_1 = E f_1, \quad H f_2 = E f_2 \quad (01)$$

donde H es el operador hamiltoniano

$$H = (1/2m) \mathbf{P}^2 + V \quad (02)$$

siendo P el operador (vectorial) de momentum lineal y V el operador (escalar, multiplicativo) de potencial.

De otro lado, sabemos que si un electrón penetra en un campo magnético \mathbf{B} el momento lineal pasa a ser $\mathbf{P} + e \mathbf{A}$, donde $\nabla \times \mathbf{A} = \mathbf{B}$ y el hamiltoniano toma la forma

$$H_B = (1/2m) (\mathbf{P} + e \mathbf{A})^2 + V \quad (02a)$$

A continuación escribiremos las dos ecuaciones escalares (01) como una sola ecuación (matricial).

$$H^\circ f = E f \quad (03)$$

donde H° es una matriz diagonal trivial, f es una función vectorial y E es un número real

$$H^\circ = \begin{vmatrix} H & 0 \\ 0 & H \end{vmatrix} = H I, \quad f = \begin{vmatrix} f_1 \\ f_2 \end{vmatrix}$$

siendo I la matriz unidad. Notemos que también podemos escribir

$$H^\circ = (1/2m) \mathbf{P}^2 I + V I, \quad (02b)$$

Es claro que (03) es una forma trivial de reescribir las dos ecuaciones (01), las mismas que, además, son coincidentes.

Ahora, teniendo presente que \mathbf{P} es un operador (vectorial) de tres componentes, $\mathbf{P} = h/i \nabla$, construiremos un operador $\hat{\mathbf{P}}$ multiplicando cada componente por un coeficiente matricial,

$$\hat{\mathbf{P}} = \hat{u}_1 P_1 + \hat{u}_2 P_2 + \hat{u}_3 P_3 \quad (04)$$

donde los \hat{u}_k son ciertas matrices de dos por dos. Nótese que no interesa el orden en el que se escriban los productos $\hat{u}_k P_k$. Las matrices \hat{u}_k deben ser construídas de manera que se cumpla

$$\hat{\mathbf{P}}^2 = \mathbf{P}^2 I \quad (04a)$$

es decir,

$$H^\circ = (1/2m) \hat{\mathbf{P}}^2 + VI \quad (02c)$$

Elevando al cuadrado la expresión (04) y teniendo presente que las componentes de \mathbf{P} conmutan entre sí, obtenemos

$$\begin{aligned} \hat{\mathbf{P}}^2 = & \hat{u}_1^2 P_1^2 + \hat{u}_2^2 P_2^2 + \hat{u}_3^2 P_3^2 + (\hat{u}_1 u_2 + \hat{u}_2 \hat{u}_1) P_1 P_2 \\ & + (\hat{u}_2 u_3 + \hat{u}_3 \hat{u}_2) P_2 P_3 + (\hat{u}_3 u_1 + \hat{u}_1 \hat{u}_3) P_3 P_1 \end{aligned}$$

Ahora, con el objeto de satisfacer la condición (04a) exigimos que las matrices \hat{u}_k satisfagan las condiciones

$$\hat{u}_1^2 = \hat{u}_2^2 = \hat{u}_3^2 = I \quad (04b)$$

$$\hat{u}_k \hat{u}_j + \hat{u}_j \hat{u}_k = 0 \quad \text{para } k \neq j \quad (04c)$$

Hasta este momento, la ecuación (03), con el hamiltoniano (02c) sólo es una forma trivial de escribir la misma ecuación de Schroedinger para un electrón en ausencia de un campo magnético. Ahora conectaremos un campo magnético B , lo que, según (02a) y (02c), nos lleva a considerar el hamiltoniano (matricial)

$$H^{\circ}_B = (1/2m) (P + e A)^2 + V I \quad (02d)$$

Este hamiltoniano ya no es diagonal, pues si bien las componentes del operador P conmutan entre sí y las componentes del operador (multiplicativo) A también conmutan entre sí; en cambio, las componentes de P no conmutan con las componentes de A , por lo cual las componentes del operador $P + e A$ tampoco conmutarán entre sí. La correspondiente ecuación

$$H^{\circ}_B f = E^{\circ}_B f \quad (03a)$$

donde E°_B es un número real, representa a dos ecuaciones (escalares) acopladas.

Ahora, teniendo presente tanto las relaciones (04b), (04c) como

$$\begin{aligned} P_i P_j - P_j P_i &= 0, & A_i A_j - A_j A_i &= 0; \text{ y también} \\ P_i A_j - A_j P_i &= (h/i) \partial A_j / \partial x_i, \end{aligned}$$

podemos desarrollar $(P + e A)^2$ de la siguiente manera:

$$\begin{aligned} (P + e A)^2 &= (P + e A)^2 + e (P_1 A_2 + A_1 P_2 - P_2 A_1 - A_2 P_1) \hat{u}_1 \hat{u}_2 \\ &\quad + e (P_2 A_3 + A_2 P_3 - P_3 A_2 - A_3 P_2) \hat{u}_2 \hat{u}_3 \\ &\quad + e (P_3 A_1 + A_3 P_1 - P_1 A_3 - A_1 P_3) \hat{u}_3 \hat{u}_1 \\ &= (P + e A)^2 + (he/i) (\partial A_2 / \partial x_1 - \partial A_1 / \partial x_2) \hat{u}_1 \hat{u}_2 \\ &\quad + (he/i) (\partial A_2 / \partial x_3 - \partial A_3 / \partial x_2) \hat{u}_2 \hat{u}_3 \\ &\quad + (he/i) (\partial A_3 / \partial x_1 - \partial A_1 / \partial x_3) \hat{u}_3 \hat{u}_1 \\ &= (P + e A)^2 + (he/i) (B_3 \hat{u}_1 \hat{u}_2 + B_1 \hat{u}_2 \hat{u}_3 + B_2 \hat{u}_3 \hat{u}_1) \end{aligned}$$

es decir,

$$(P + e A)^2 = (P + e A)^2 I + 2e S_B \quad (05)$$

donde

$$S_B = (h/2i) (B_1 \hat{u}_2 \hat{u}_3 + B_2 \hat{u}_3 \hat{u}_1 + B_3 \hat{u}_1 \hat{u}_2) \quad (05a)$$

lo cual nos permite escribir

$$H_B = H_B I + (e/m) S_B \quad (02e)$$

donde, como deseábamos, en el nuevo hamiltoniano aparece un término adicional, dependiente del campo magnético.

Ahora veamos la forma que debe tener S_B . De (04b) y (04c) obtenemos que $\hat{u}_k = -\hat{u}_1 \hat{u}_k \hat{u}_1$, de donde resulta que la traza de cada una de tales matrices debe ser nula,

$$\text{Tr}(\hat{u}_k) = 0, \quad k = 1, 2, 3$$

es decir, dichas matrices deben tener la forma,

$$\begin{vmatrix} q & r \\ s & -q \end{vmatrix}$$

Si tenemos presente que las matrices buscadas deben representar a un observable físico, ellas deberán ser autoadjuntas, es decir, el número q debe ser real y, además, debe cumplirse que $s = r^*$. Por otra parte, debido a (04b), los elementos de la matriz deben satisfacer que $q^2 + rs = 1$. La forma más general de una matriz que satisface los requerimientos mencionados es:

$$\begin{vmatrix} \cos \beta & e^{i\mu} \sin \beta \\ e^{-i\mu} \sin \beta & -\cos \beta \end{vmatrix} \quad (06)$$

donde μ, β son números reales arbitrarios. Ahora, puesto que para un sistema de tres matrices autoadjuntas, siempre podemos encontrar una transformación unitaria que diagonalice a una de ellas, entonces para \hat{u}_3 podemos tomar el valor $\beta = 0$, obteniendo

$$\hat{u}_3 = \begin{vmatrix} 1 & 0 \\ 0 & -1 \end{vmatrix} \quad (06a)$$

Si, ahora, de acuerdo con (04c), exigimos que las matrices de la forma (06) deban anticonmutar con la matriz (06a), obtenemos $\cos \beta = 0$, es decir $\beta = \pi/2$. Entonces, las otras dos matrices tendrán la forma

$$\hat{u}_1 = \begin{vmatrix} 0 & e^{i\mu_1} \\ e^{-i\mu_1} & 0 \end{vmatrix} \quad \hat{u}_2 = \begin{vmatrix} 0 & -e^{i\mu_2} \\ e^{-i\mu_2} & 0 \end{vmatrix}$$

pero, por otra parte, estas matrices deben anticonmutar entre sí, $\hat{u}_1 \hat{u}_2 + \hat{u}_2 \hat{u}_1 = 0$. De esta condición resulta que, necesariamente $\cos(\mu_1 - \mu_2) = 0$, es decir $\mu_1 - \mu_2 = \pi/2$. Ahora, por comodidad escribimos $\mu_1 = \mu$; entonces obtenemos $e^{i\mu_2} = ie^{i\mu}$, con lo cual podemos escribir

$$\hat{u}_1 = \begin{vmatrix} 0 & e^{i\mu} \\ e^{-i\mu} & 0 \end{vmatrix} \quad \hat{u}_2 = i \begin{vmatrix} 0 & -e^{i\mu} \\ e^{-i\mu} & 0 \end{vmatrix} \quad (06b)$$

donde μ es un número real arbitrario. Para el caso particular en el que $\mu = 0$, se obtienen las conocidas matrices de Pauli.

Usando (06a) y (06b) podemos verificar que, al multiplicar tales matrices, se cumplen las relaciones

$$\hat{u}_j \hat{u}_k = i \hat{u}_m \text{ para } j, k, m = 1, 2, 3 \text{ cíclicamente}$$

lo cual, con (05a), nos permite escribir

$$S_B = (h/2) (\hat{u}_1 B_1 + \hat{u}_2 B_2 + \hat{u}_3 B_3) = (h/2) \hat{\mathbf{u}} \cdot \mathbf{B} \quad (05b)$$

Si ahora definimos las matrices

$$S_k = (h/2) \hat{u}_k, \quad \mathbf{S} = (S_1, S_2, S_3) \quad (07)$$

podemos verificar que se satisfacen las relaciones

$$S_j S_k = (ih/2) S_m, \text{ para } j, k, m = 1, 2, 3 \text{ cíclicamente,}$$

$$S_j S_k + S_k S_j = 0, \text{ para } j \neq k$$

de las cuales podemos obtener

$$\mathbf{S} \times \mathbf{S} = i\hbar \mathbf{S}$$

que es la misma ecuación que satisfacen los operadores de momento angular.

Usando (05b), (06a), (06b) y (7) obtenemos

$$S_B = (h/2) \begin{vmatrix} B_3 & e^{i\mu} B_- \\ e^{-i\mu} B_+ & -B_3 \end{vmatrix} \quad (05c)$$

donde $B_- = B_1 - i B_2$, $B_+ = B_1 + i B_2 = B_-^*$, lo que evidencia que el hamiltoniano H^B , en contraste con el operador H^0 , no es trivialmente diagonal.

También podemos verificar que las columnas

$$\tilde{a}_+ = \begin{vmatrix} B_+ & B_3 \\ e^{-i\mu} & B_- \end{vmatrix} \quad \tilde{a}_- = \begin{vmatrix} e^{i\mu} & B_- \\ -B_+ & B_3 \end{vmatrix} \quad (08)$$

donde B es el módulo de \mathbf{B} , son los vectores propios del operador S_B ,

son los vectores propios del operador S_B ,

$$S_B \tilde{a}_\uparrow = (+Bh/2) \tilde{a}_\uparrow, \quad S_B \tilde{a}_\downarrow = (-Bh/2) \tilde{a}_\downarrow \quad (08a)$$

Los vectores \tilde{a}_\uparrow and \tilde{a}_\downarrow son ortogonales pero no unitarios. Para obtener que ellos sean unitarios escribiremos el campo B en coordenadas esféricas expresadas en algún sistema de referencia (arbitrario),

$$B_1 = B \sin\phi \cos\beta, \quad B_2 = B \sin\phi \sin\beta, \quad B_3 = B \cos\phi$$

lo cual permite escribir

$$B_+ = B e^{i\beta} \sin\phi = 2B \cos(\phi/2) e^{i\beta} \sin(\phi/2)$$

$$B + B_3 = 2B \cos(\phi/2) \cos(\phi/2),$$

construir los vectores propios ortonormales

$$\tilde{A}_\uparrow(\phi, \beta) = \begin{vmatrix} e^{i(\mu-\beta)} \cos(\phi/2) \\ \sin(\phi/2) \end{vmatrix}, \quad \tilde{A}_\downarrow(\phi, \beta) = \begin{vmatrix} e^{i(\mu-\beta)} \sin(\phi/2) \\ -\cos(\phi/2) \end{vmatrix} \quad (08b)$$

y reescribir al operador S_B

$$S_B = S \cdot B(\phi, \beta) = B h/2 \begin{vmatrix} \cos\phi & e^{i(\mu-\beta)} \sin\phi \\ e^{-i(\mu-\beta)} \sin\phi & -\cos\phi \end{vmatrix} \quad (05d)$$

Energías y funciones propias del operador H°_B

Si definimos la constante numérica

$$b = (eh/2m) B$$

y tenemos presente las expresiones (02e) y (05d), podemos escribir (03a) en la forma

$$\begin{vmatrix} H_B + b \cos\phi & b e^{i(\mu-\beta)} \sin\phi \\ b e^{-i(\mu-\beta)} \sin\phi & H_B - b \cos\phi \end{vmatrix} f = E^\circ_B f \quad (3b)$$

o, si se prefiere,

$$[H_B + b \cos\phi] f_1 + [b e^{i(\mu-\beta)} \sin\phi] f_2 = E^\circ_B f_1 \quad (3c)$$

$$[b e^{-i(\mu-\beta)} \sin\phi] f_1 + [H_B - b \cos\phi] f_2 = E^\circ_B f_2 \quad (3d)$$

donde eventualmente debemos considerar dos valores para E°_B .

Sean, por otra parte g_1 y g_2 dos funciones propias del operador H_B

$$H_B g_1 = E_B g_1, \quad H_B g_2 = E_B g_2 \quad (9)$$

correspondientes a un mismo valor propio (no exigimos que dichas funciones sean linealmente independientes). Ahora mostraremos que tales funciones son soluciones de las ecuaciones (3c) y (3d). En efecto, aplicando (9) a (3c) y (3d) podemos escribir

$$[E_B + b \cos\phi] g_1 + [b e^{i(\mu-\beta)} \sin\phi] g_2 = E^\circ_B g_1$$

y

$$[b e^{-i(\mu-\beta)} \sin\phi] g_1 + [E_B - b \cos\phi] g_2 = E^\circ_B g_2$$

cuyas soluciones no triviales exigen que los coeficientes de las funciones consideradas satisfagan la ecuación

$$[E_B + b \cos\phi - E^\circ_B] [E_B - b \cos\phi - E^\circ_B] - b^2 \sin^2\phi = 0$$

la misma que proporciona dos valores para la energía E°_B

$$E^\circ_{B1} = E_B + b, \quad E^\circ_{B2} = E_B - b$$

o también

$$E^\circ_{B1} = E_B + (eh/2m) B, \quad E^\circ_{B2} = E_B - (eh/2m) B \quad (10)$$

que son los valores de la energía que se obtienen en el experimento de Stern-Gerlach. Además obtenemos que las correspondientes funciones (vectoriales) propias son

$$f_\uparrow = \begin{vmatrix} (e^{i(\mu-\beta)} \cos\phi/2) g_1 \\ (\sin\phi/2) g_1 \end{vmatrix} = A_\uparrow(\phi, \beta) g_1, \quad f_\downarrow = A_\downarrow(\phi, \beta) g_1$$

donde g_1 es una de las funciones de (9).

Conclusión

Partiendo de la ecuación no relativística de Schroedinger, $H f_1 = E f_1$, o de la ecuación $H_B f_1 = E f_1$ cuando se haya presente un campo magnético B , hemos construido la ecuación matricial $H^\circ f = E^\circ f$, o correspondientemente, la ecuación $H^\circ_B f = E^\circ_B f$, donde $H^\circ_B = H_B I + (e/m) S_B$.

Cuando no existe un campo magnético la ecuación $H^\circ_B f = E^\circ_B f$ se reduce a expresar dos veces la ecuación $H f_1 = E f_1$. Pero en la presencia de un campo magnético, el hamiltoniano H_B es modificado por la adición de un sumando, que explica correctamente los dos niveles de energía medidos después que el rayo de electrones emerge del aparato de Stern-Gerlach, (10).

Finalmente queremos anotar que:

- i) Los valores propios de S_B son completamente independientes de la dirección del campo B . Esto es consistente con el hecho de que tal dirección está referida a un sistema de referencia arbitrario.
- ii) Los vectores propios de S_B son vectores en un espacio complejo y no representan ninguna dirección en el espacio físico convencional.
- iii) Los vectores propios "rotan" al rotar el campo B , con un período 4π , de manera que

$$\bar{A}_\uparrow(\phi + 4\pi, \beta + 4\pi) = \bar{A}_\uparrow(\phi, \beta), \quad \bar{A}_\downarrow(\phi + 4\pi, \beta + 4\pi) = \bar{A}_\downarrow(\phi, \beta)$$

pero

$$\bar{A}_\uparrow(\phi + 2\pi, \beta + 2\pi) = -\bar{A}_\uparrow(\phi, \beta), \quad \bar{A}_\downarrow(\phi + 2\pi, \beta + 2\pi) = -\bar{A}_\downarrow(\phi, \beta)$$

y

$$\bar{A}_\uparrow(\phi + 3\pi, \beta) = \bar{A}_\downarrow(\phi, \beta), \quad \bar{A}_\downarrow(\phi + 3\pi, \beta) = \bar{A}_\uparrow(\phi, \beta)$$

- iv) En el caso especial en el que el tercer eje del sistema de referencia coincida con la dirección de B , es decir, para $\phi = 0$, el operador S_B se torna diagonal y las ecuaciones (3c) y (3d) se desacoplan, dando

$$[H_B + (eh/2m) B] f_1 = E_{B1} f_1, \quad [H_B - (eh/2m) B] f_2 = E_{B2} f_2$$

- v) Los ángulos ϕ y β expresan la orientación (arbitraria) del campo B . En ninguna parte se ha mencionado una supuesta orientación que podría atribuírsele al spin del electrón.
- vi) En ningún momento hemos hecho uso de la inhomogenidad del campo B ; sólo nos ha interesado la presencia del mismo.

REFERENCIAS

- [1] H. Valqui. *Expansión de un operador por efecto de un prisma*; *Rev. Per. de Física*, Vol. 1, N° 2, 1981.
- [2] S. Fluegge. *Lehrbuch der Theoretischen Physik, Band 4*; Springer 1964.
- [3] D. Blochinzew. *Grundlagen der Quantenmechanik*; Harri Deutsch 1962.
- [4] Feynman-Leighton-Sands. *The Feynman Lectures on Physics, III*; Addison-Wesley 1965.

Tesis

Diseño y construcción de un microscopio de efecto túnel

Fernando Camino (*)

Desde los primeros exitosos experimentos de G. Binnig, H. Rohrer y colaboradores, la microscopía de efecto túnel se ha convertido en una técnica invaluable para el análisis de superficies e interfases. En este trabajo se detallan los pasos del diseño y construcción de un microscopio de efecto túnel de pequeño tamaño ($<1\text{ cm}^3$), masa pequeña y con un método de aproximación gruesa basado en el deslizamiento inercial. El cabezal del MET tiene una alta frecuencia de resonancia producto de la pequeña masa del cabezal y también de la independencia del ambiente, mediante un control electrónico autónomo que reduce considerablemente el esfuerzo en aislar mecánicamente el sistema. Se explican los detalles del diseño de la electrónica de control, la construcción del cabezal del microscopio así como el software necesario para la operación del mismo.

(*) Facultad de Ciencias, Universidad Nacional de Ingeniería, Casilla postal 31-139, Lima - Perú.

Introducción

Breve descripción del funcionamiento del MET

En 1981 Binnig, Rohrer y colegas¹ introdujeron el Microscopio de Efecto Túnel (MET). El principio del MET es simple: una punta muy fina se aproxima muy cerca de la muestra por medio de un sistema de realimentación, mediante el barrido de la punta sobre la muestra (similar al barrido que efectúan los electrones en una pantalla de TV) se obtiene una imagen a partir de la señal de realimentación (fig. 1). La punta así como la muestra son conductoras, y mediante la aplicación de un pequeño voltaje entre la punta y la muestra hacemos que una corriente de tunelaje empiece a fluir, considerando que la distancia entre la punta y la muestra es muy pequeña (del orden de un nanómetro). Esta corriente es fuertemente dependiente de la distancia y midiendo la corriente, la distancia puede ser determinada. La punta se mueve mediante piezoeléctricos cerámicos, los cuales cambian sus dimensiones al aplicárseles voltaje. La resolución del MET es muy alta, tiene resolución atómica y aquí reside su gran importancia, i.e. cerca de 1 Å lateralmente y 0.01 Å verticalmente.

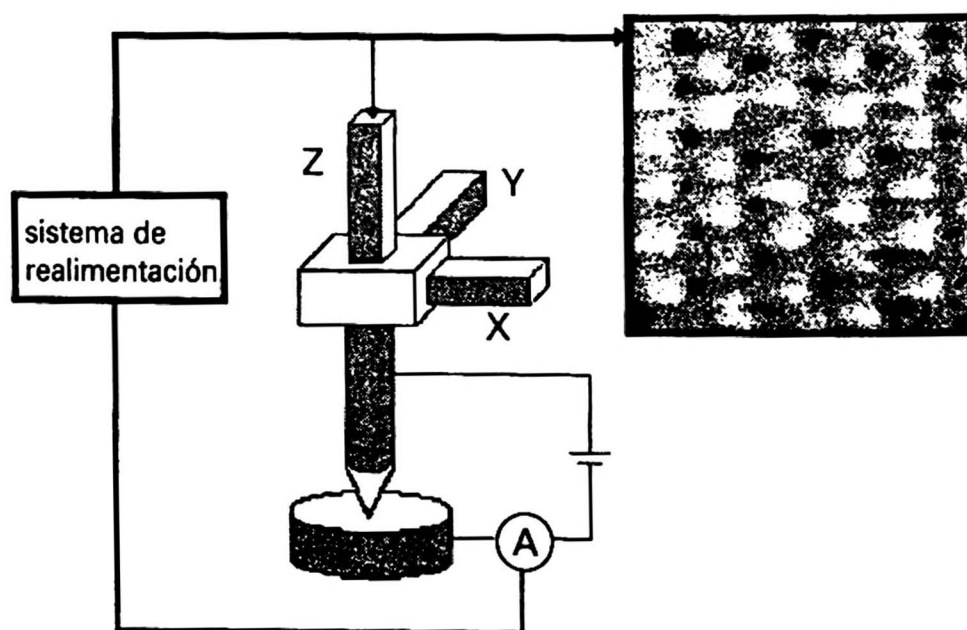


Fig. 1 En el MET la imagen se forma haciendo un barrido de una punta muy afilada sobre la muestra. Un sistema de realimentación mantiene la distancia punta-muestra (alrededor de 10 Å) constante controlando el voltaje aplicado al piezoelemento z. La distancia se determina midiendo la corriente túnel entre la punta conductora y la muestra conductora. Esta corriente es fuertemente dependiente de la distancia. Cuando se hace que la punta haga un barrido sobre la muestra aplicando voltajes apropiados a los piezos-x e y, el sistema de realimentación ordena al piezo-z que mueva la punta de tal forma que siga la superficie a una distancia punta-muestra constante. La señal de realimentación provee información de la topografía de la superficie y graficando esta señal, generamos una imagen. La imagen en la figura es una imagen real de un monocristal de grafito, tomada con nuestro MET. La distancia entre puntos claros (átomos de la primera etapa atómica superficial) es de 2,46 Å.

Principio del tunelaje

Existen dos modos de abordar teóricamente el efecto del tunelaje en el MET. El primero se conoce como el principio de dispersión, donde la ecuación de Schrödinger se resuelve acoplando una onda incidente con una onda reflejada y transmitida. Este modelo es simple, sin embargo, el cálculo de la probabilidad de transmisión puede ser muy difícil. Simmons² solucionó este problema en una dimensión utilizando el método WKB y un metal con electrones libres, lo que es un modelo muy simplificado. El segundo método se le conoce como el del hamiltoniano de transferencia³, en donde la punta y la muestra se consideran dos sistemas independientes con una interacción débil. Ambas teorías han sido utilizadas con éxito en uniones planas donde la distancia de separación es de 10 nm y el tunelaje es esencialmente unidimensional debido a la fuerte dependencia de la distancia. El gap de túnel en el MET es mucho más pequeño, cerca de 1 nm, y la punta se supone aguda y no plana. Esto hace al método del hamiltoniano de transferencia dudoso pues está limitado a una interacción débil entre los electrodos. Sin embargo es ampliamente utilizado ya que permite la incorporación de muchas partículas y efectos inelásticos en geometrías tridimensionales. A continuación daré una breve explicación del método de dispersión el cual aunque bastante aproximado es suficiente para justificar la dependencia exponencial de la corriente de tunelaje respecto a la distancia punta-muestra.

En la solución dada por Simmons², para ondas planas y un metal de electrones libres, la corriente de tunelaje I se puede describir en función del ancho de la barrera z y el voltaje aplicado V , por

$$I \propto \frac{\phi}{z^2} \exp(-A \phi^{\frac{1}{2}} z) - \frac{\phi + eV}{z^2} \exp[-A (\phi + eV)^{\frac{1}{2}} z], \quad (1)$$

donde ϕ es la altura media de la barrera por encima del nivel de Fermi del electrodo a potencial negativo, y la constante $A = (2m)^{1/2}/h$ (m es la masa del electrón y h la constante de Planck). La relación es también válida para altos voltajes en el modo de emisión de campo. En el límite de pequeños voltajes, la relación de corriente puede ser simplificada a

$$I \propto \frac{V}{z} \exp(-A \phi^{\frac{1}{2}} z) \quad (2)$$

La principal propiedad de la ec. (2) es la dependencia exponencial con la distancia punta-muestra z . Un cambio en distancia de tunelaje de 1 Å con $\phi = 4$ eV, cambiará la corriente de tunelaje en un orden de magnitud.

Construcción

Introducción

Para conseguir resolución atómica, el aislamiento de vibraciones es esencial. Existen dos formas de conseguir una solución adecuada. La primera es hacer la unidad del MET (el cabezal) tan rígido como sea posible. El segundo es reducir la transmisión de la vibración ambiental. Un sistema de supresión de vibraciones comúnmente usado consiste en un conjunto de resortes de suspensión y un mecanismo de amortiguamiento.

Todo cabezal del MET debe contener un mecanismo grueso de aproximación que permita reducir la distancia punta-muestra hasta la distancia de tunelaje. Se desea que este mecanismo sea lo más pequeño posible de manera que no disminuya la rigidez del cabezal. La elección del mecanismo de aproximación gruesa determina en gran medida la performance del cabezal final del MET.

Los sistemas de aproximación manual basados en motores a pasos, utilizando tornillos, resortes y palancas⁴ son simples pero no son soluciones compactas. Estos sistemas son susceptibles a influencias de la temperatura y la vibración. Existen otras formas de extender el rango de los piezoeléctricos. Las más comunes son: 1. multiplicación mediante pasos estáticos como en el "louse"⁵ o el inchworm,^{6,7} y 2. movimiento inercial,⁸ que utiliza la inercia de la parte a mover contra rápidos movimientos del piezoelemento. Nosotros utilizamos el sistema introducido por S. Kleindiek y K.H. Herrmann⁹, modificado ligeramente. Este es un posicionador tridimensional con aproximación fina y gruesa de ajuste puramente electrónico. Este sistema mejora la estabilidad mecánica mediante una rigurosa reducción del tamaño y la masa del sistema de posicionamiento, resultando en una alta frecuencia mecánica de resonancia que garantiza una eficiente protección contra las vibraciones ambientales. La idea básica es reducir el MET a unas pocas partes esenciales: Dos tubos concéntricos posicionadores para el movimiento lateral y longitudinal de la punta conectados por una base, un pedazo de alambre de Pt/Ir con la punta, un porta-muestra soportado por el piezotubo x/y. Para desacoplar el sistema mecánicamente del ambiente, las líneas de alimentación eléctricas deben de ser tan finas como sea posible. El sistema utiliza un mecanismo similar al que utiliza la Ref. 7, salvo que aquí lo utilizamos para controlar la aproximación de la punta a la distancia túnel.

El cabezal

La pieza principal del MET la constituye sin lugar a dudas el llamado cabezal del MET. El cabezal es la pieza en donde se ubican los piezoeléctricos de barrido, el sistema de aproximación fina y gruesa, la punta y la muestra.

El diseño de nuestro cabezal depende principalmente del método de aproximación gruesa que hemos elegido. El MET utiliza el método de deslizamiento inercial. En la fig. 2 podemos apreciar un esquema simplificado del arreglo mecánico

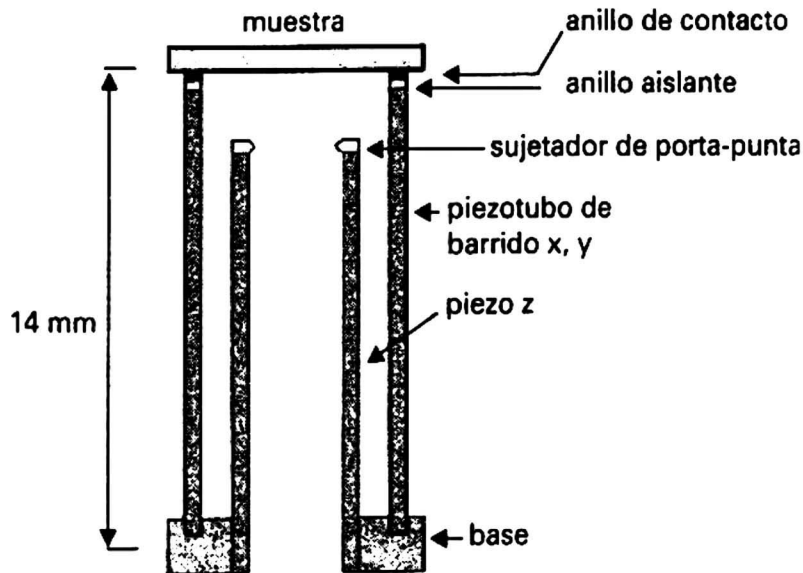


Fig. 2 Diseño esquemático del cabezal del MET. No se muestra el porta-punta. Ambos piezoeléctricos cilíndricos están montados en una base cerámica hecha a partir de lava volcánica.

Como se aprecia en la fig. 2, se montan dos piezotubos concéntricos en una base cerámica. El piezotubo externo tiene 12,7 mm de largo; 6,35 mm de diámetro y 0,5 mm de espesor. Este piezotubo tiene su electrodo externo dividido en cuatro sectores y es el encargado de desplazar la muestra lateralmente respecto de la punta (que se ubica en el eje de los piezotubos). El piezotubo interno mide 12,7 mm de largo; 3,2 mm de diámetro y 0,5 mm de espesor, y se encarga exclusivamente del desplazamiento vertical de la punta. El material de los piezos es PZT-5H y sus electrodos son de níquel.

La primera frecuencia de resonancia longitudinal del piezo z es de 55 kHz y la frecuencia de resonancia en el modo de barrido es de 22 kHz. Valores muy aceptables.

El porta-punta se aprecia en la fig. 3 (a) y en la fig. 3(b) se aprecia el conjunto completo.

Sistema de aproximación gruesa

Como hemos mencionado líneas arriba, el método de aproximación gruesa escogido es el de deslizamiento inercial. La velocidad de acercamiento se puede regular,

(70 nm/s como mínimo para una amplitud del diente de 60 V). La velocidad típica de aproximación es de alrededor de 360 nm/s, aunque es bastante lenta (se demora 23 min en avanzar 1 mm) es aconsejable ya que una vez llegada a la zona túnel no se quiere que la punta siga avanzando por inercia.

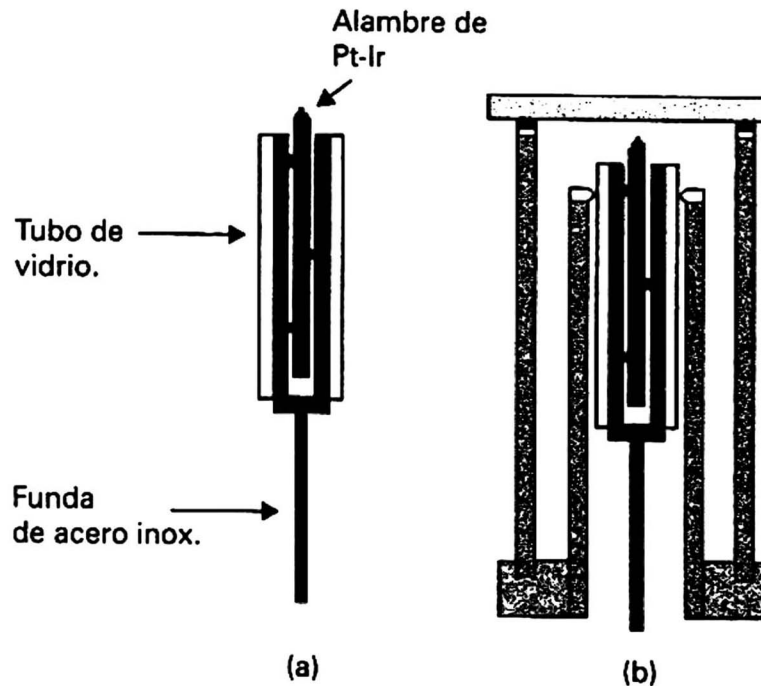


Fig. 3 (a) Porta-punta del MET. (b) Cabezal completo. Al aplicar un voltaje de diente de sierra al piezo z haremos que el porta-punta (y por consiguiente la punta de Pt-Ir) se mueva hacia la muestra por fricción y deslizamiento inercial.

Con 360 nm/s, lo que la punta avanza por inercia una vez llegada a la zona túnel (considerando una razón fricción/masa de 100 m/s^2) es $6,5 \times 10^{-6} \text{ Å}$, es decir, nada. Con un periodo del diente de 0,8 ms, la velocidad de aproximación es de 0,45 mm/s y lo que avanza la punta por inercia es 1 nm, es decir, la distancia túnel promedio, pudiendo con mucha probabilidad, chocar la punta con la muestra. También se desea una velocidad de aproximación muy lenta, con el fin de considerar el tiempo de respuesta de la etapa electrónica. Una vez llegada la punta a la zona túnel, la electrónica no detiene inmediatamente el diente de sierra, sino que demora un tiempo muy corto pero finito, luego una velocidad muy lenta, asegura que la punta no avance demasiado durante este lapso de tiempo.

Idealmente, podemos suponer que la punta se mantiene en su posición durante el tiempo de caída del diente, por lo tanto en cada ciclo del diente, la punta avanza una distancia de 360 nm hasta que la electrónica detecte la corriente túnel y detenga la punta. En la fig. 4 podemos apreciar el movimiento de la punta respecto del movimiento del piezo z.

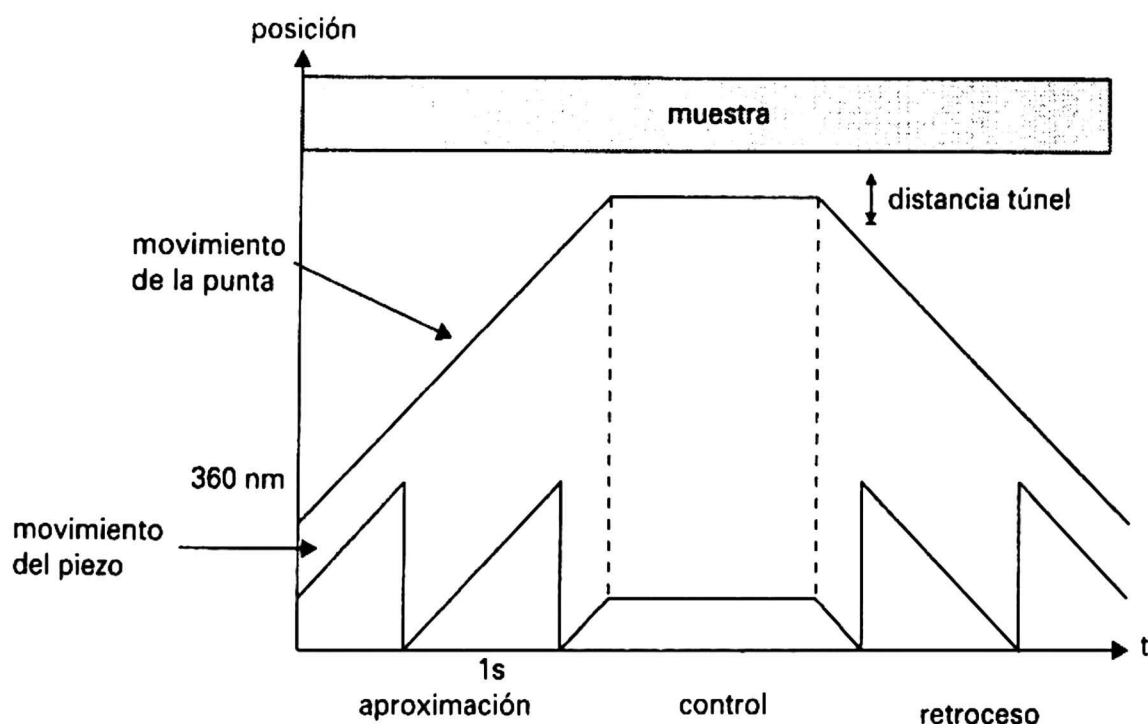


Fig. 4 La electrónica entrega el diente de sierra hasta que se detecte la corriente túnel. En la zona de control se procede a la medición, aquí el movimiento de la punta es muy fino y la aguja y piezo se mueven en conjunto (sin deslizamiento). Una vez terminada la medición, la electrónica entrega un diente de sierra de rampa inversa alejando la aguja de la muestra.

Movimiento fino

La aproximación de la punta cesa cuando se detecta la corriente túnel. De aquí se procede a tomar una imagen de la muestra. Al barrer el área de la muestra, el piezo z mueve la punta muy levemente, con el fin de mantener una corriente túnel constante durante el barrido. La punta se desplaza $60 \text{ \AA}/V$. Por lo tanto si se aplican voltajes de $0,1 \text{ mV}$, podemos desplazar la punta distancias por debajo de $0,01 \text{ \AA}$. Aunque en principio, la resolución vertical del MET puede fácilmente estar por debajo de $0,01 \text{ \AA}$, por problemas de estabilidad mecánica, no es posible ir por debajo de este nivel.

La electrónica

La etapa electrónica del MET tiene tres partes principales: el amplificador de corriente, el sistema de realimentación y los amplificadores que controlan el voltaje aplicado al piezotubo de barrido horizontal (ejes x , y).

El amplificador de corriente se construye alrededor de un amplificador operacional de alta impedancia y muy baja corriente "bias" (por ejemplo, el OPA128 de la Burr-Brown). Usualmente la corriente túnel está en el rango de $1\text{-}10 \text{ nA}$, y esta corriente

se detecta haciendo que el amplificador de voltaje detecte la caída de voltaje a través de un resistor de 1-100 M Ω . Lo más conveniente es colocar el amplificador lo más cerca posible de la unión túnel para minimizar la capacitancia.

Los componentes del sistema de realimentación se muestran en la Fig. 5. Usualmente la corriente túnel se linealiza mediante un amplificador logarítmico para compensar la dependencia exponencial de la corriente túnel con la distancia. Esta señal se compara con una señal de referencia seguida de un regulador proporcional integral el cual controla el voltaje aplicado al piezo z (el piezoeléctrico encargado del movimiento vertical de la punta). Muchos diseños utilizan solamente la parte integral del control PI.

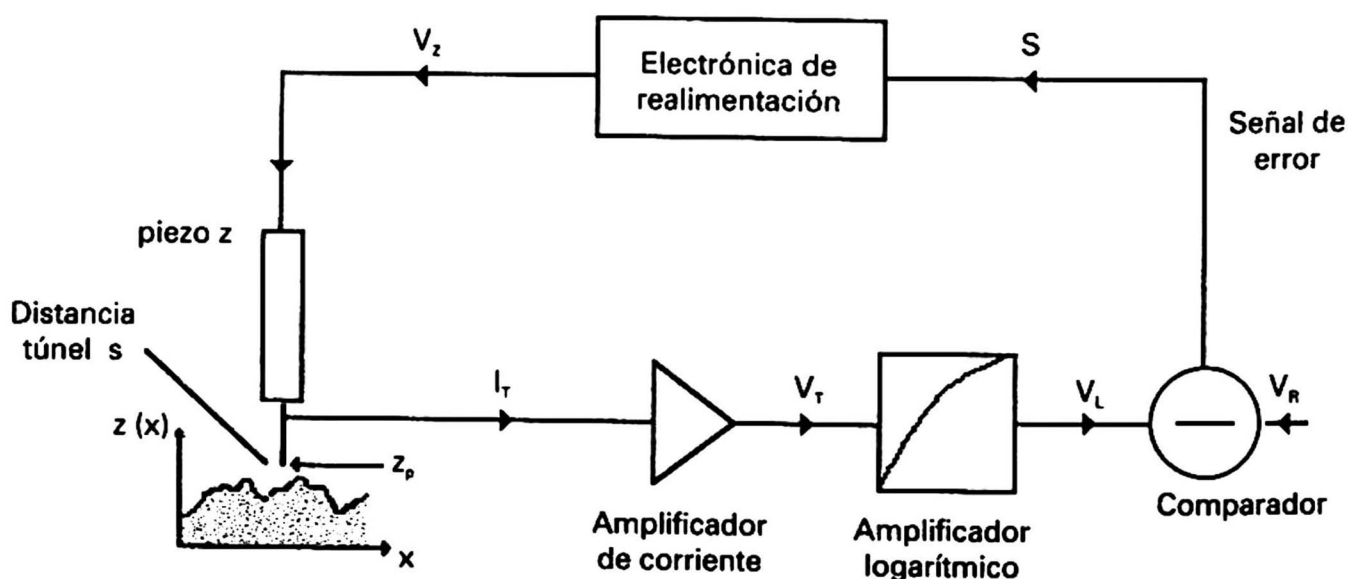


Fig. 5 Diagrama esquemático del lazo de realimentación en el MET.

Se desea que la ganancia total del sistema de realimentación sea lo más alta posible sin causar oscilaciones.

Nuestro diseño hace que los amplificadores que manejan al piezotubo de barrido sean amplificadores operacionales convencionales que operan a ± 18 V, evitando el uso de costosos amplificadores de alto voltaje. Estos amplificadores permiten seleccionar diferentes rangos de barrido, cambiando los valores de sus ganancias. Reciben las señales de barrido provenientes de los conversores D/A y se lo entregan a los electrodos adecuados del piezotubo de barrido.

Circuito de control

El sistema de control consta de dos etapas: el integrador y los Schmitt triggers.

El integrador.- El integrador es la etapa estándar de control. La fig. 6 muestra su diseño. La función de esta etapa es la de mantener la distancia punta-muestra constante durante una medición V_t es la señal túnel proveniente del preamplificador y V_{ref} define la corriente túnel que queremos mantener constante.

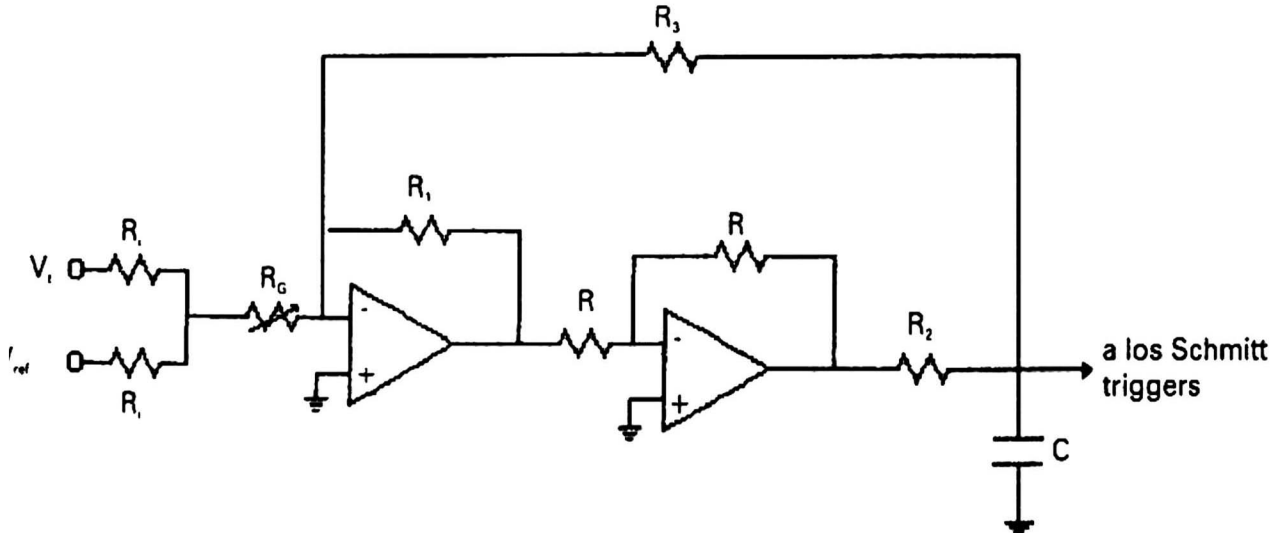


Fig. 6 Integrador. V_{ref} define la corriente túnel y R_G (que es variable) ajusta el ancho de banda del lazo y la ganancia de la etapa.

Si cumplimos con la condición $R_1 = R_2 + R_3$, el voltaje V_c en el condensador es

$$V_c = (V_t + V_{ref}) \frac{R_1}{R_2 (R_i + 2 R_G) C} t \quad (3)$$

Los valores que tenemos son $R_1 = 1 \text{ M}\Omega$, $R_2 = 150 \text{ k}\Omega$, $R_G = 0-1 \text{ M}\Omega$, y $C = 470 \text{ pF}$. Luego la constante de esta etapa varía entre 6500 y 94500 $\text{V}^{-1}\text{s}^{-1}$. Altos valores significan respuestas rápidas pero con baja estabilidad. La salida de esta etapa es la señal de retroalimentación que se le aplica al piezo z para mantener la distancia punta-muestra constante. Esta es la señal que nos da la información de la topografía de la muestra.

Los Schmitt triggers.- Cuando añadimos el circuito de la fig. 7 a la etapa de integración, tenemos el sistema que permite generar los dientes de sierra de aproximación (o alejamiento) hasta llegar a la zona túnel en forma automática.

Cuando la punta está alejada de la muestra, la rampa del diente tendrá una pendiente que depende del voltaje de referencia y de la ganancia del integrador, tal como lo muestra (3). Podemos llegar a velocidades de aproximación muy lentas definiendo un V_{ref} pequeño. La ec. (3) nos dice que la rampa aumentará de valor

linealmente hasta que se llegue al voltaje de corte definido por el Schmitt-trigger (comparador con histéresis) ST2. Una vez llegado a este valor, los diodos conducen y el condensador se descarga a través de ellos, comenzando una rampa positiva nuevamente. En el retroceso de la punta el proceso es similar salvo que la descarga se realiza por los diodos asociados con ST1.

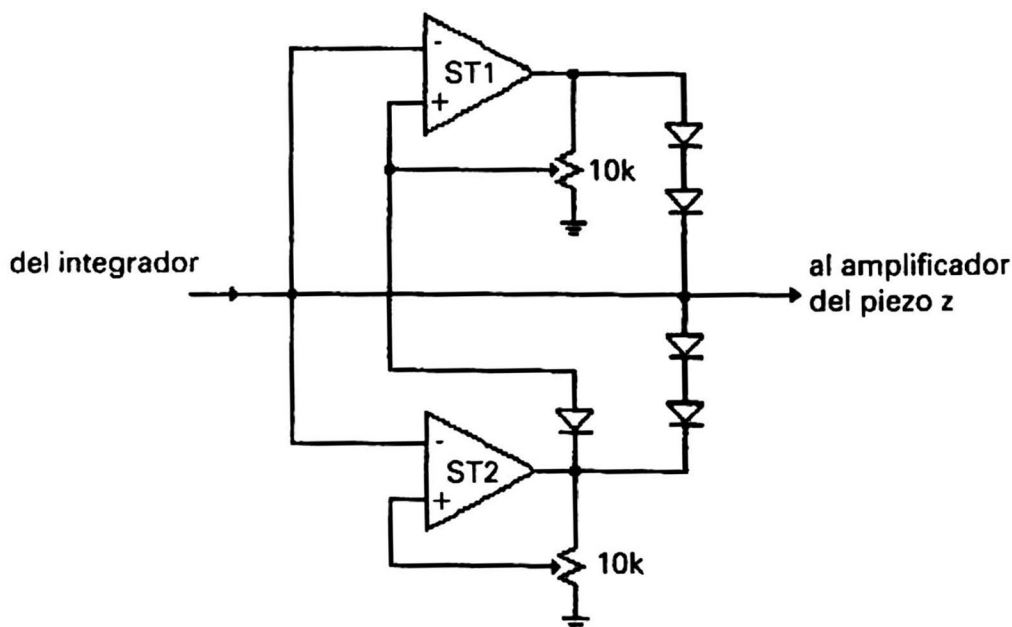


Fig. 7 Schmitt triggers. Los diodos son 1N4148. ST2 corta la rampa positiva (acercamiento) y ST1 corta la rampa negativa (alejamiento).

Computadora

Casi todos los microscopios de efecto túnel de hoy en día hacen uso de una computadora. La computadora puede ser utilizada tanto para las medidas como para la presentación de las imágenes. La tarjeta de interfase entre la computadora y la electrónica del MET necesita al menos dos conversores D/A para comandar los desplazamientos laterales (x, y) y un conversor A/D para almacenar el valor del voltaje aplicado al piezo z el cual lleva la información de la topografía de la muestra. Es preferible que estos conversores sean de 16 bits de resolución.

El software del MET debe de ser capaz de mostrar las imágenes después o durante un barrido. Lo puede hacer como una imagen en escala de grises (en dos dimensiones, el tono de gris vendría a darnos el dato de altura). Existen programas más elaborados permiten definir la localización de la fuente luminosa, creando efectos de sombras que resaltan las propiedades de la muestra. La calidad de la imagen puede ser ampliada por muchas técnicas. Una muy común es crear un plano de referencia desde el cual se miden las distancias, lo que permite corregir imágenes que están un poco inclinadas. Se puede reducir el ruido de la medición mediante filtros digitales. Además todo software de MET debe de tener la facilidad de almacenar las imágenes que se tomen.

Resultados experimentales

A continuación mostraremos imágenes con resolución atómica de la superficie de grafito pirolítico altamente orientado (HOPG). Las imágenes fueron tomadas en el modo de corriente constante. El grafito es un material popular en los estudios de microscopía de efecto túnel, ya que es muy fácil preparar una superficie limpia con grandes superficies atómicamente planas simplemente partiendo (cleaving) la muestra. El grafito es un material hecho por capas, donde las capas se mantienen unidas por fuerzas débiles de Van der Waals, lo que explica la facilidad de "clivelar" el material (el procedimiento práctico es adherir una cinta adhesiva a una cara del cristal y luego despegarla con cuidado, en la cinta quedará unida una capa muy fina dejando una superficie limpia y de gran área en el cristal).

Dentro de una capa, los átomos están ubicados en un arreglo hexagonal formado por fuertes enlaces π entre los átomos. En la fig. 8 se muestra una imagen obtenida con nuestro MET.

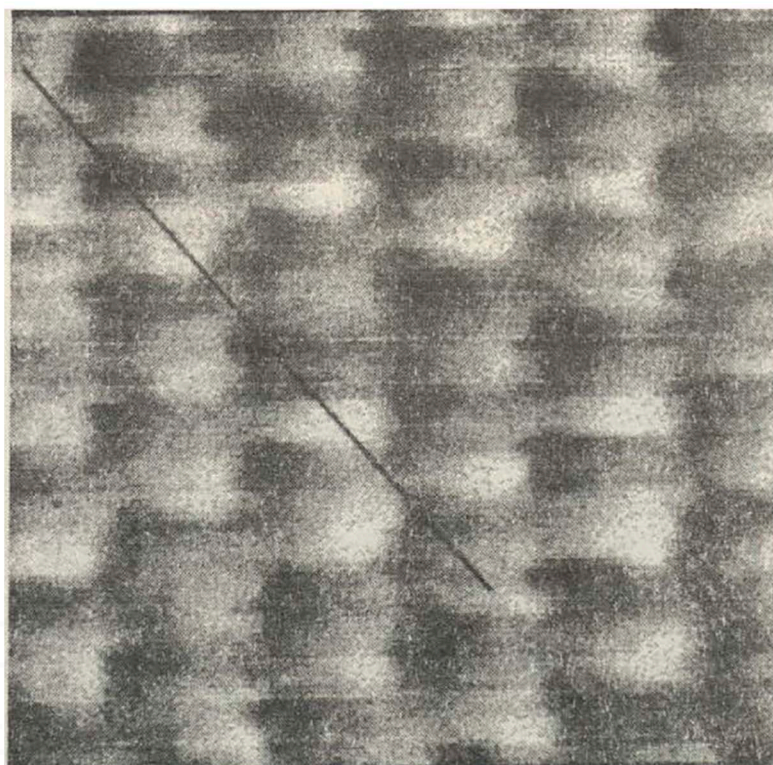


Fig. 8 Imagen suavizada vista de arriba de HPOG. $I = 0,5$ nA, V punta-muestra = $0,1$ V, voltaje de barrido = $0,1$ V, número de columnas = número de filas = 128, promedio = 20. Con punta cortada de Pt-Ir. Los átomos visibles son solamente los ubicados en las posiciones β . La recta representa las posiciones por las que se ha efectuado un corte transversal, como se explicará más adelante.

La imagen de la fig. 8 no muestra la estructura hexagonal del grafito más bien muestra una simetría triangular. La interacción capa-capa puede explicar este hecho. Las posiciones de los átomos de carbón en una capa no son equivalentes (fig. 9). La

posición α tiene un átomo de carbón directamente debajo en la segunda capa, más no la posición β . Los enlaces π en las posiciones α interactúan con la capa inferior, conduciendo a nuevos enlaces bastante alejados del nivel de Fermi. Como hemos visto anteriormente, a bajos voltajes punta-muestra, el MET sólo muestra los estados en el nivel de Fermi, y consecuentemente una imagen con el MET sólo debe de mostrar las posiciones β . No sólo observamos esta anomalía con las imágenes de grafito. Existen otras anomalías tales como corrugaciones gigantes de la superficie de grafito, que son estudiadas y nos dan nuevas informaciones acerca de las propiedades elásticas de la superficie, tales como potenciales interatómicas y constantes elásticas locales.

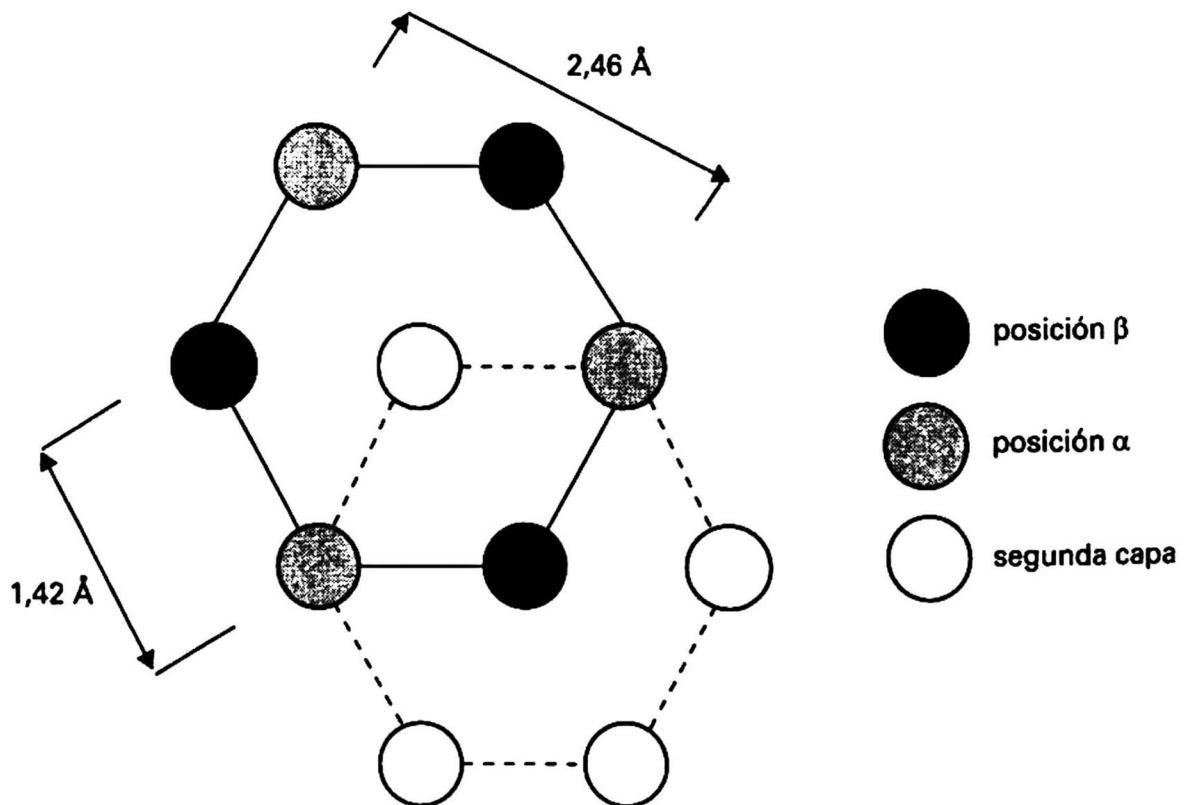


Fig. 9 Dibujo esquemático de dos capas de grafito. Las posiciones α , tiene un átomo directamente debajo en la segunda capa, mientras que las posiciones β no.

Según la fig. 9 los átomos negros (posiciones β) serán los que más sobresalen en la superficie, y por lo tanto se presentarán blancos en la imagen. Las posiciones α no conducen suficiente corriente túnel como para notarse claramente en la imagen, aunque la corriente que conducen no es completamente nula. Esto es justamente lo que se muestra en la imagen de la fig. 8, donde se aprecia claramente la estructura de triángulo equilátero que forman los átomos vecinos en las posiciones β .

La facilidad de corte transversal, nos puede confirmar lo arriba dicho. Si realizamos un corte transversal a lo largo de la recta que parte de una posición β y

biseca el lado opuesto, el corte cruzará sucesivamente una posición β , un hueco (átomo de la segunda capa) y una posición α , en forma periódica. El gráfico del corte transversal, debe de presentar picos de corriente para las posiciones β , mínimos en los huecos y pequeños picos para las posiciones α , ya que como hemos mencionado anteriormente, la corriente túnel en estas posiciones aunque pequeña, no es completamente nula. Esto lo apreciamos en la fig. 10, que muestra el resultado del corte transversal según la recta que se muestra en la imagen de la fig. 8.

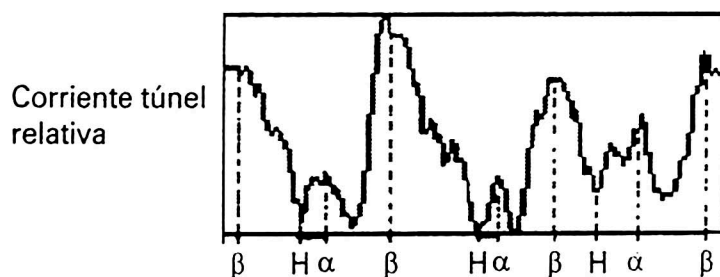


Fig. 10 Corte transversal según la recta de la imagen de la fig. 8. Esta recta pasa por cuatro máximos que corresponden a átomos en las posiciones β . H significa un hueco, es decir, una posición de un átomo de la segunda capa, sin ningún átomo sobre él.

Para calibrar el desplazamiento lateral, utilizamos la imagen de la fig. 11.

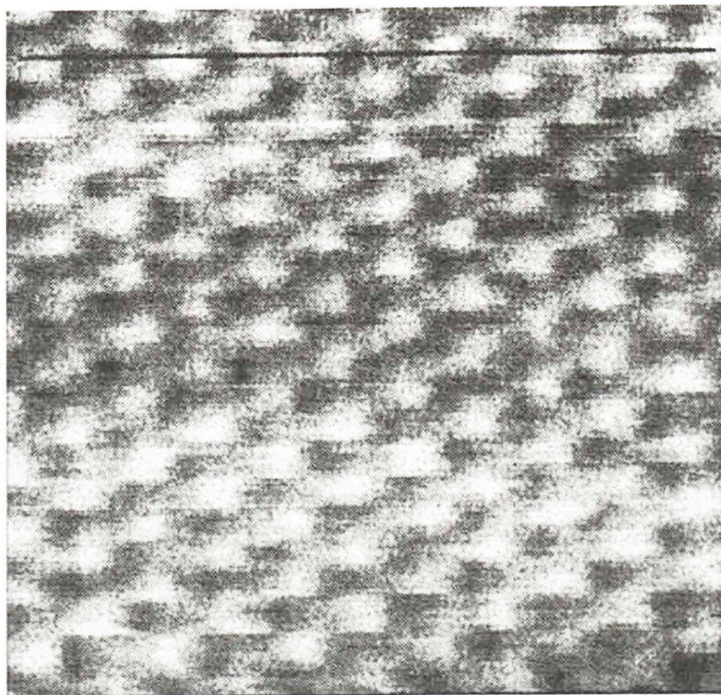


Fig. 11 Imagen suavizada vista de arriba de HPOG. $I=0,5$ nA; V punta-muestra = 0,1; voltaje de barrido = 0,205 V; número de columnas = número de filas = 128, promedio = 20. Con punta cortada de Pt-Ir.

Si analizamos la línea trazada en la fig. 11, vemos que pasa por 6 posiciones β . Dos posiciones β consecutivas forman un rombo a partir de dos triángulos equiláteros unidos por su base, tal como se muestra en la fig. 12.

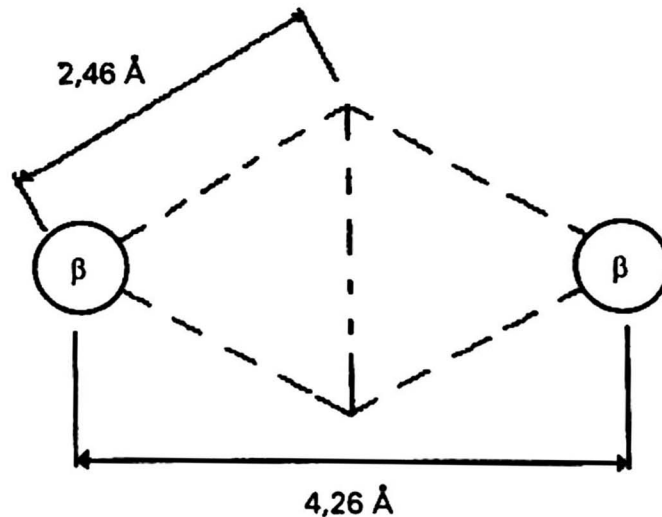


Fig. 12 La distancia entre dos posiciones β a lo largo de la recta trazada en la imagen de la fig. 11.

En el ancho de la muestra, según la recta trazada, existen cinco de estos rombos. luego el ancho de la muestra es de $5 \times 4,26 = 21,3 \text{ Å}$. Como el barrido lateral se hizo con un voltaje de $0,205 \text{ V}$, tenemos que la calibración en los ejes laterales es de $103,9 \text{ Å/V}$. Este valor experimental, está muy próximo al valor de 125 Å/V calculado teóricamente. La diferencia se debe a las propiedades intrínsecas de los piezoelementos cerámicos como el "arrastre" (creep) e histéresis durante el barrido.

En la fig. 13 apreciamos una imagen tridimensional de la superficie mostrada en la fig. 8.

No he detallado en este informe partes importantes del MET tales como el aislamiento de vibraciones, el software y la construcción de las puntas. Una explicación detallada de la construcción del MET está disponible en mi tesis de licenciatura.

Conclusiones

- En esta tesis se presentó el trabajo realizado para construir en el Perú el primer microscopio de efecto túnel con resolución atómica en los aspectos mecánico, electrónico y de computación. Se mostraron las primeras imágenes de grafito pirolítico obtenidas en atmósfera de aire, a temperatura ambiente.

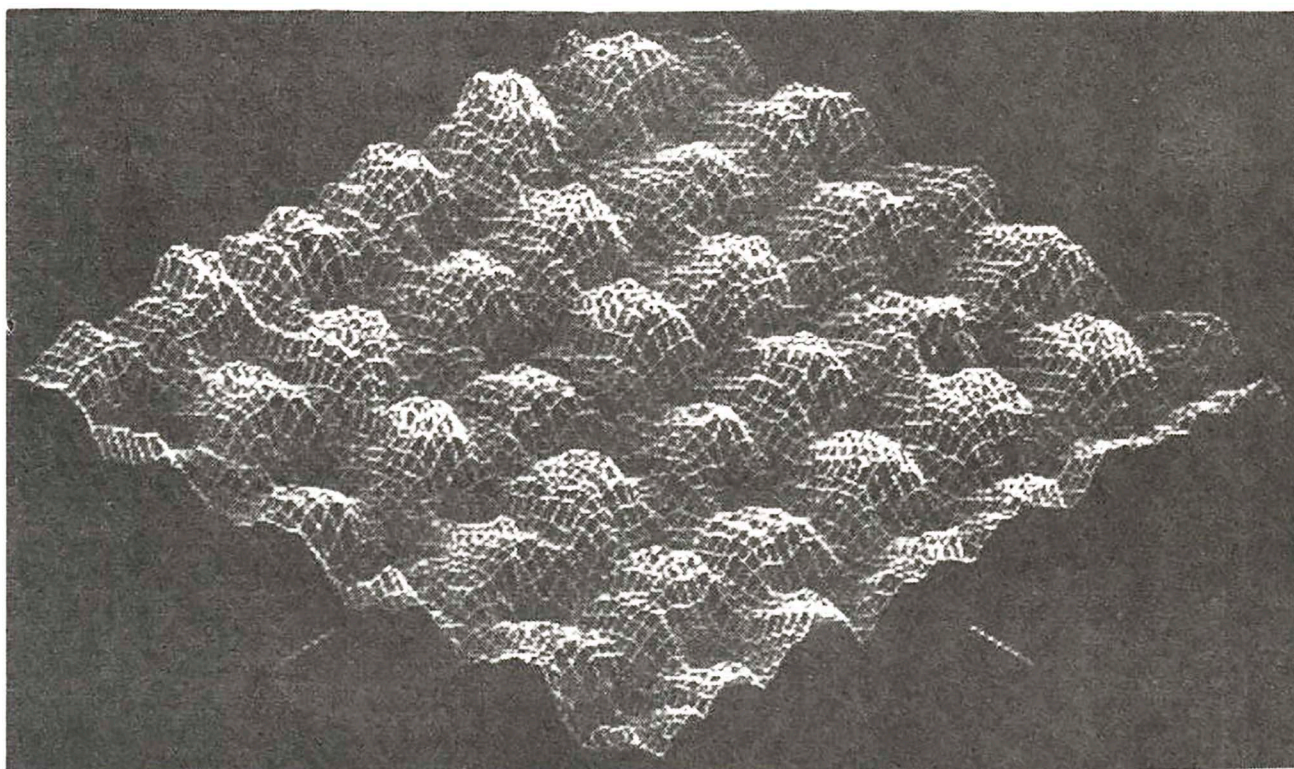


Fig. 13 Imagen tridimensional de la superficie mostrada en la fig. 8. Interlineado (densidad de la malla) = 1, ángulo azimutal = 50° , rotación en el plano horizontal, 45° .

- El sistema presenta una excelente supresión de vibración acústica. La resonancia lateral del MET se estima alrededor de 20 kHz, y la vertical en 100 kHz. Buena compensación térmica, y alta velocidad de operación. Tan altos valores en las frecuencias de resonancia son producto de la pequeña masa del cabezal y también de la independencia del ambiente, mediante un control electrónico autónomo que reduce considerablemente el esfuerzo en aislar mecánicamente el sistema.
- El principal propósito que llevó al Prof. Estrada a la construcción del MET fue la de familiarizarnos con él, para luego aplicarlo en el estudio del crecimiento de películas delgadas sobre superficies sólidas para lo cual tendremos que diseñar un sistema que opere en medio líquido. En un futuro cercano pensamos incursionar en el fascinante campo de las nanoestructuras.

Agradecimientos

Primero, quiero agradecer a mi profesor asesor Dr. Walter Estrada por darme la oportunidad de participar en este fascinante proyecto. Su confianza y apoyo fueron indispensables para culminar esta etapa del trabajo.

Deseo agradecer al Dr. Boris Chornik del Departamento de Física de la Universidad de Chile quien nos apoyó en todo momento y nos permitió aprender todo lo que sabemos acerca del Microscopio de Efecto Túnel.

Agradezco a mis co-trabajadores, el Licenciado Ciro Carhuacho quien diseñó y construyó las tarjetas electrónicas del Microscopio y a Grabierl Altuna.

BIBLIOGRAFÍA

- ¹ Binnig, H. Rohrer, C. Gerber, y E. Weibel, *Phys. Rev. Lett.* 49, 57 (1982).
- ² G. Simmons, *J. Appl. Phys.* 34, 1793 (1963).
- ³ J. Bardeen, *Phys. Rev. Lett.* 6, 57 (1961).
- ⁴ F. J. Giessibl, Ch. Gerber, y G. Binnig, *J. Vac. Sci. Technol. B* 9, 984 (1991).
- ⁵ G. Binnig, H. Rohrer, Ch. Gerber y E. Weibel, *Appl. Phys. Lett.* 40, 178 (1982).
- ⁶ *Inchworm* es una marca registrada de Burleigh Instruments.
- ⁷ D. R. Peale y B. H. Cooper, *J. Vac. Sci. Technol. A* 8, 345 (1990).
- ⁸ K. Besocke, *Surf. Sci.* 181, 145 (1987).
- ⁹ S. Kleindiek y K. H. Herrmann, *Rev. Sci. Instrum.* 64, 693 (1993).

Estudio electrocrómico, estructural y morfológico de películas delgadas de óxido de níquel preparadas por la técnica del rociado pirolítico

Mónica Marcela Gómez León (*)

RESUMEN

Películas de óxido de níquel, con propiedades electrocrómicas remarcables, fueron obtenidas usando spray-pyrolysis. En un trabajo previo, usando la misma técnica, se obtuvo óxido de níquel electrocrómico a partir de una solución acuosa de nitrato de níquel; sin embargo la transmitancia óptica en el estado claro fue menor del 60% (para 630 nm. de longitud de onda), básicamente debido a la rugosidad de la superficie, haciendo este material aún no apto para aplicaciones ópticas. En este trabajo se presenta los resultados obtenidos para las películas de óxido de níquel producidas por rocío pirolítico a partir de una solución alcohólica de nitrato de níquel. Las películas muestran una transmitancia óptica en el estado claro (para 630 nm. de longitud de onda), mayor de 85% y un contraste óptico entre el estado claro y el estado coloreado de aproximadamente 50%, indicando un buen comportamiento electrocrómico. Los recubrimientos o películas obtenidas de la solución alcohólica muestran una estructura NiO-bunsenita.

(*) Facultad de Ciencias, Universidad Nacional de Ingeniería, Lima, Casilla postal 31-139, Lima - Perú.

Introducción

Para obtener películas delgadas de óxido de níquel dirigidas hacia las aplicaciones electrocrómicas se han empleado varias técnicas, siendo la mayoría de ellas basadas en técnicas de vacío [1,2] o por procedimientos electroquímicos [3,4]. Sin embargo, la técnica de spray pyrolysis ha sido escasamente explotada [5], a pesar de ser ésta una técnica sencilla, de bajo costo y potencialmente útil para la fabricación de recubrimientos en áreas grandes.

En un trabajo previo [5], nosotros reportamos el comportamiento electrocrómico de las películas de óxido de níquel obtenidas por spray pirólisis utilizando una solución acuosa de nitrato de níquel como la solución pirolítica. Estas películas presentaban un importante electrocromismo, sin embargo para aplicaciones prácticas ("smart windows"), el estado claro de las películas mostraban poca transmitancia óptica (<60%). Nosotros asociamos este efecto a la rugosidad de la superficie de la película, intrínseca a las condiciones de depósito en la técnica empleada.

En este trabajo nosotros presentamos las características electrocrómicas del óxido de níquel obtenidas variando las condiciones pirolíticas y teniendo como base una solución alcohólica diluida de nitrato de níquel. Bajo condiciones apropiadas, se ha logrado obtener una mejora substancial de las propiedades ópticas del material respecto a aquellas obtenidas de soluciones acuosas.

Experimento

Las películas de óxido de níquel fueron obtenidas utilizando un sistema de rocío pirolítico tal como se describió en un trabajo previo [5], utilizando aire como gas portador del aerosol. El comportamiento electrocrómico fue estudiado usando un potencióstato WENKING POS 73 en una celda electrocrómica con tres electrodos (usando Calomel Saturado como electrodo de referencia) para la voltametría cíclica (inserción/extracción de carga), y un espectrofotómetro (OPTOMETRIC, RS 25) para la medida de la transmitancia óptica espectral. Para la voltametría cíclica se utilizó como electrolito una solución acuosa de KOH 0,1 M.

Las películas electrocrómicas fueron depositadas en un sustrato de vidrio previamente recubierto con ITO (película conductora transparente). Las películas optimizadas fueron obtenidas a la temperatura de 300°C. El espesor de las películas fueron medidas por el método interferométrico óptico. Las soluciones alcohólicas de nitrato de níquel empleadas para el rocío fueron soluciones alcohólicas de nitrato de níquel 0,05; 0,1 y 0,25M.

El análisis estructural fue llevado a cabo en un difractómetro de rayos X.

Resultados y discusión

La superficie de las muestras obtenidas a partir de soluciones alcohólicas de nitrato de níquel muestran una menor rugosidad que las obtenidas en soluciones acuosas. Este efecto fue claramente observado por microscopía óptica. Estos resultados indican que en una solución alcohólica el solvente no interfiere en toda la reacción pirolítica en el sustrato para temperaturas mayores que 300 °C. Nosotros proponemos que la siguiente reacción estaría llevándose a cabo en el sustrato:



La razón de crecimiento de las películas fue proporcional al contenido de la sal en la solución rociada; sin embargo para concentraciones mayores que 0,25 M el soluto empieza a precipitar. Por tanto las películas optimizadas fueron obtenidas a esta concentración.

Las medidas de difracción de rayos X muestran que las películas tienen una estructura cúbica del NiO-bunsenita (fig. 1). Después del tratamiento electroquímico se observa que los picos de difracción se incrementan notablemente en el estado claro indicando que la cristalinidad aumenta después de la inserción de carga dentro de la película electrocrómica.

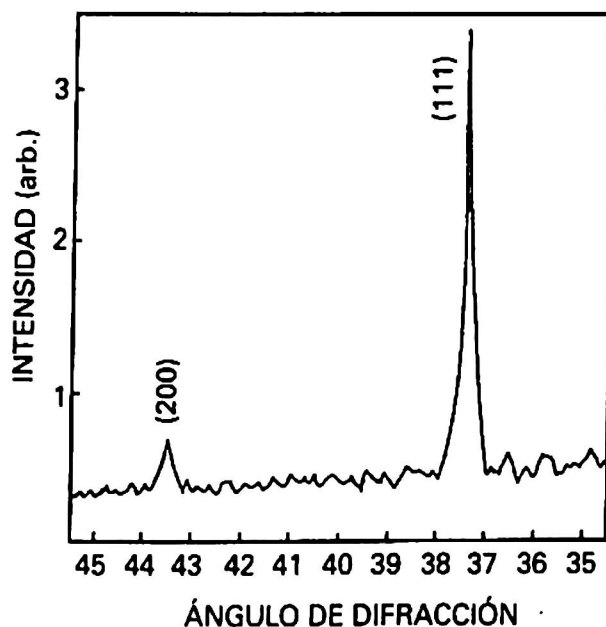


Fig. 1 Patrón de difracción de rayos-X para una muestra de NiOx producida a 400°C. Los picos corresponden a la estructura cúbica de NiO.

El efecto electrocrómico evoluciona conforme se va efectuando el ciclaje voltamétrico. Se puede observar en la fig. 2 que el área del perfil voltamétrico va incrementándose conforme va aumentando el número de ciclos. En el ciclo 1 apenas

se observa el pico anódico; este va incrementándose hasta obtenerse un perfil voltamétrico estable (ciclo 20). Después de 20 ciclos tanto los perfiles de voltametría como de transmitancia óptica (fig. 2) no experimenta cambios mayores.

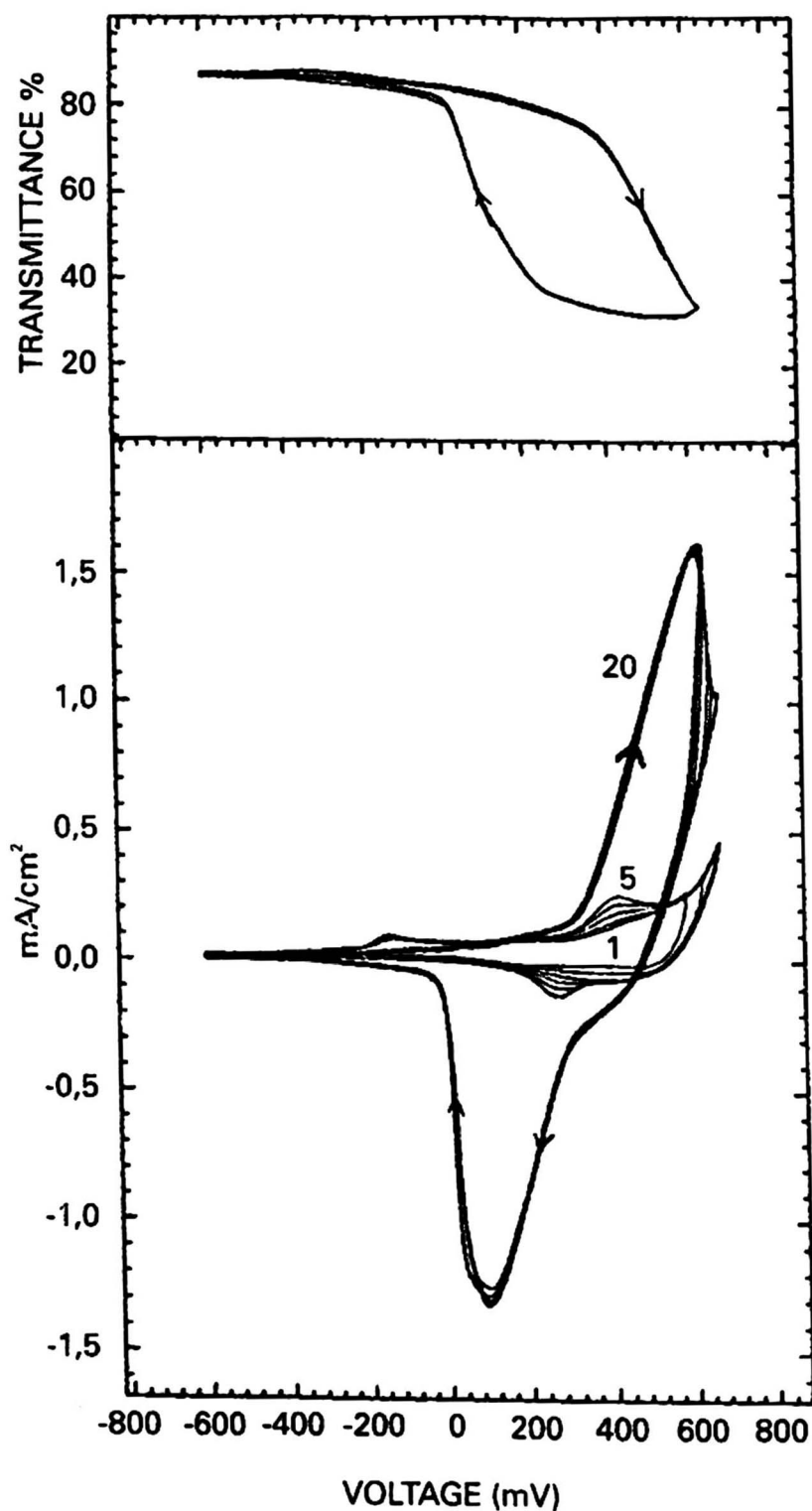


Fig. 2 Transmitancia a 630 nm in-situ con la voltametría cíclica para una muestra de óxido de níquel a 1 ciclo, 5 ciclos y 20 ciclos.

La fig. 3 muestra las medidas de la transmitancia espectral tanto para el estado claro (0 mC/cm^2 de carga extraída) como para el estado coloreado (30 mC/cm^2 de carga extraída), para una película de espesor de $0,5 \mu\text{m}$. El electrocromismo es pronunciado en la región $400\text{-}600 \text{ nm}$; por ejemplo en 550 nm se alcanza una variación entre los estados claro y oscuro del 50%; así mismo la transmitancia para el estado claro fue del 90%. Este espectro mostró poca variación después de 2000 ciclos voltamétricos.

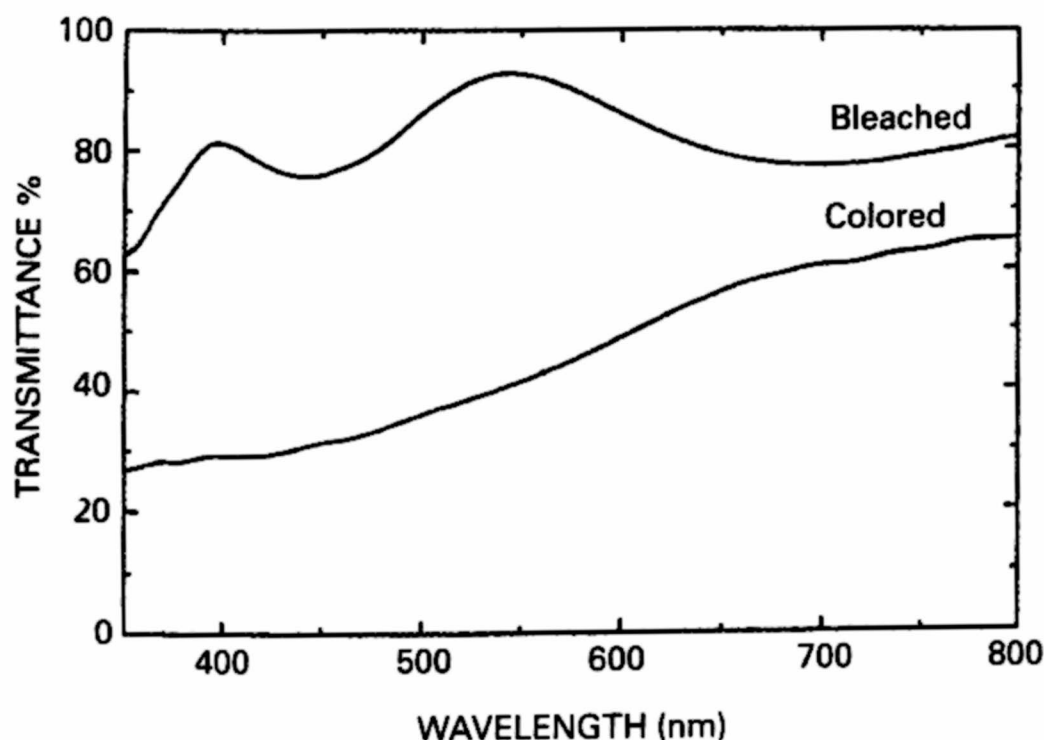


Fig. 3 Transmitancia espectral de las películas electrocromáticas de NiO, en el estado transparente y coloreado.

Este resultado indica claramente que las películas de óxido de níquel pirolíticas, producidas a partir de una solución alcohólica, tienen un buen comportamiento electrocromático, comparable a aquellas obtenidas por técnicas de vacío [1]. La baja rugosidad superficial observada, respecto a aquellas obtenidas por solución acuosa [5], sería la razón por la cual estas películas presentan una alta transmitancia (más del 80%) en el estado transparente.

Conclusiones

Un buen comportamiento electrocromático en las películas de óxido de níquel puede conseguirse utilizando una solución alcohólica de nitrato de níquel, mostrando un contraste óptico y durabilidad comparable a aquellas obtenidas por técnicas de vacío.

Este resultado indica que usando spray-pyrolysis, una técnica sencilla y de bajo costo, es posible obtener óxido de níquel electrocrómico con propiedades similares a aquellas obtenidas por técnicas más sofisticadas y de mayor costo.

REFERENCIAS:

1. J. S.E.M. Svensson and C. G. Granqvist, *Sol. Energy Mater.* 16 (1987) 19.
2. W. Estrada, A. Andersson and C. G. Granqvist, *J. Appl Phys.* 64 (1988) 1802.
3. M. Fantini and A. Gorestein, *Sol. Energy Mater.* 16 (1987) 487.
4. M. Carpenter, R.S. Conell and D.A. Corrigan, *Sol. Energy Mater.* 16 (1987) 333.
5. J. Arakaki, R. Reyes, M. Horn and W. Estrada, *Sol. Energy Mater.* 37 (1995) 33.

Análisis de excursiones de potencia por inserciones de reactividad en reactores de investigación

Gerardo Lázaro Moreyra (*)

RESUMEN

Se estudia la conducta del núcleo de un reactor tipo piscina cuando éste es sometido a excursiones de potencia por inserciones violentas de reactividad. Se inserta reactividad de \$1,5 a \$3,2 en tiempos de 0,10 s a 0,80 s. La excursión se aplica a dos núcleos, el A16/21 y el A21/25 y los resultados son comparados con los SPERT-I: B24/32, B16/64 y D12/25. La potencia al inicio del accidente se varía de 5 watts a 10 MW y la refrigeración del núcleo, de convección natural a convección forzada en 1650 m³/h. Se realiza cálculos de importancia de algunas variables de entrada del código en cuanto a la magnitud del pico de potencia y la energía liberada. Estos datos pueden ser definidos en variables de tres grupos: a) En los datos cinéticos se elije el beta efectivo por ser muy sensible a la configuración de trabajo. b) En transferencia de calor se elije varias opciones de correlaciones de

*) Facultad de Ciencias / Post Grado, Universidad Nacional de Ingeniería, Instituto Peruano de Energía Nuclear - Centro Nuclear de Huarangal / RP10.

predicción de regímenes de transferencias de calor y c) En la ingeniería de seguridad se varía el tiempo de retardo para el inicio efectivo de caída de barras, este tiempo demuestra ser transcendente para la integridad del núcleo, aún cuando los primeros y últimos tramos de las barras de seguridad inducen una pobre reactividad con respecto a su zona central. Esto permite que la reactividad del sistema se mantenga negativa y no ocurra una nueva excursión o en todo caso este segundo pico no se desarrolle plenamente tal como sucede con $\beta = 1,5/800$ ms y 200 ms de retardo para el núcleo A16/25. Este tiempo de retardo, por su importancia, se determina experimentalmente. También se determina los coeficientes de reactividad por generación de vacío y por cambio de temperatura del moderador; el tiempo total de inserción de barras y el peso en reactividad de las barras de seguridad y control para los dos núcleos A16/21 y A16/25.

Se investiga la evolución de los parámetros de autocontrol del núcleo en cada uno de los mecanismos de compensación: cambio de densidad de moderador, cambio de temperatura de moderador, cambio de temperatura de combustible, y por dilatación de los elementos combustibles. Se cuantifica todas ellas en el transcurso de la evolución del transitorio y se observa graficados su importancia relativa. Se determina también la máxima reactividad que puede introducirse al núcleo sin alcanzar temperaturas de fusión del revestimiento de la placa combustible.

Introducción

Para la simulación se utilizó el programa de cálculo PARET [1, 2, 3]. La versión usada es el paquete implementado para una PC [4], y se aplica para condiciones iniciales de potencia de 5 watts a 10 MW, con inserciones de reactividad de $\beta = 1,5$ a $\beta = 3,0$ en 800 ms.

Para todos los casos analizados, el sistema de Seguridad Banco de barras de seguridad entra a mitigar el accidente sólo cuando el flujo de potencia supere el valor de los 12 MW. Se considera que la única señal que genera la lógica de seguridad para la caída de barras, por cuanto las señales termohidráulicas son muy lentas para percibir el inicio de un accidente de este tipo. Como criterio de seguridad la señal por bajo valor de período aún cuando es más rápida que la señal por alto flujo. En accidentes por pérdida de caudal, las barras de seguridad caen por bajo flujo refrigerante, y se fija en un 80% del valor nominal.

El PARET tiene dos archivos de salida el OUTPUT.RES y el OUTPUT.SUM. El primero tiene información de detalle de los resultados y el segundo los datos resumidos en cada paso de tiempo. El programa PLOTSIG.FOR, que lee y procesa el archivo OUTPUT.SUM, es una extensión del programa PLOT.FOR original [2]. Este

ha sido modificado en su mayor parte para adaptarlo a datos adicionales de salida y nueva presentación de resultados gráficos.

Para propósitos de validación de los resultados se utiliza información de los Reactores SPERT I. Este es el tipo piscina de tacho abierto a presión atmosférica, con el núcleo contenido en un tanque de 1,2 m de diámetro y 3 m de altura. Por su similitud con el reactor en estudio (RP-10), se elige tres núcleos: el SPERT I B24/32[7] el SPERT I B12/64[7] y el SPERT I de 4,02% en peso de U235. Estos son referenciados con los núcleos RP10-A16/21 y RP10-A16/25, conocidos también como núcleos 14 y 16 respectivamente. Se compara también con los datos de destrucción del núcleo SPERT I D12/25 [8].

Cálculo de la potencia total instantánea

La potencia del reactor, o es especificada como dato de entrada, o es calculada como una función del tiempo determinada a través de la solución numérica de las siguientes ecuaciones de la cinética de un reactor puntual.

$$\frac{d\Phi(t)}{dt} = \frac{[p(t) - \beta]}{\Lambda} \Phi(t) + \sum_{i=1}^6 \lambda_i C_i(t) + S_0(t)$$

$$\frac{dC_i(t)}{dt} = \frac{\beta f_i}{\Lambda} \Phi(t) - \lambda_i C_i(t); i = 1, 2, \dots,$$

Fracción y constante de decaimiento de neutrones retardados

	Grupo 1	Grupo 2	Grupo 3	Grupo 4	Grupo 5	Grupo 6
f_i	3,83842e-2	2,086214e-2	1,887324e-1	4,072227e-1	1,299396e-1	2,709960e-2
λ_i	1,27200e-2	3,17390e-2	1,16700e-1	3,12100e-1	1,4	3,85

t = Tiempo de transitorio (s)

$p(t)/\beta$ = Reactividad neta del sistema (\$)

$\Phi(t)$ = Potencia del reactor (MW)

β = Fracción efectiva de n's retardados.

Λ = Tiempo de generación de n's prontos.

$C_i(t)$ = Concentración de n's del grupo i.

λ_i = Const. decaimiento de n's grupo i.

f_i = Fracción de n's retardados grupo i.

$$\beta = \sum_{i=1}^6 \beta_i \quad f_i = \frac{\beta_i}{\beta}$$

El análisis llevado a cabo con el PARET está basado en una evaluación acoplada de los efectos nucleares, hidrodinámicos y térmicos del núcleo. El modelo del PARET consiste de un núcleo refrigerado con agua que está representado por un máximo de cuatro elementos combustibles y sus canales de refrigeración asociados. Así, el núcleo

puede ser dividido en cuatro o menos regiones, cada una con diferente generación de potencia, flujo másico de refrigeración, parámetros hidráulicos, coeficientes de realimentación, picos de potencia. Cada una es representada por una simple placa combustible, más su canal de refrigeración asociado. Un problema transitorio de excursión de potencia por inserción de reactividad es simulado especificando la reactividad positiva insertada externamente versus el tiempo de inserción.

Coeficientes de realimentación del RP10

	A16/21	A16/25
T. Moderador $^{\circ}\text{C} <\alpha_k>$	-1,097e-2	-5,470e-3
Vacío $\$/\%$ Vacío $<C_k>$	-0,3	-0,26
T. Combustible $^{\circ}\text{C} <\gamma_1>$	-2,220e-3	-2,236e-3
Beta efectivo	0,01	0,01

El programa tiene de salida cálculos detallados de transferencia de calor, cálculos de reactividades por realimentación, y cálculos hidrodinámicos que permiten la inversión del flujo refrigerante.

Las ecuaciones siguientes se utilizan para calcular qué reactividad de realimentación es la más importante en este tipo de reactores, cómo participa como parámetro de seguridad, su efecto inmediato para controlar la evolución abrupta del primer pico de potencia, y su participación en la evolución de todo el transitorio.

$$r_{\text{Rod}}^m = \frac{2}{V_{\text{Mod}} \beta_{\text{eff}}} \sum_{k=1}^2 X_{k2} \sum_{j=1}^{21} C_{j,k} (V_{\text{Rod}}^0)_{j,k} (y_c)_{j,k}^m - r_{\text{Rod}}^0$$

$$r_{\text{Rod}}^0 = \frac{2}{V_{\text{Mod}} \beta_{\text{eff}}} \sum_{k=1}^2 X_{k2} \sum_{j=1}^{21} C_{j,k} (V_{\text{Rod}}^0)_{j,k} (y_c)_{j,k}^0$$

$$V_{\text{Mod}} = \sum_{k=1}^2 X_{k2} \sum_{j=1}^{21} (V_{\text{Mod}})_{j,k}$$

$X_{k2} =$ Fracción en volumen del núcleo que representa el canal k.

$C_{j,k} =$ Coeficiente de realimentación por densidad del moderador en el incremento axial $\Delta Z_{j,k}$ ($\delta k/k/\%$ Vacío); $C_{j,k} = C_k \delta_{j,k}$, donde C_k es constante para la región k. $\delta_{j,k}$ toma en cuenta la dependencia axial.

$(y_c)_{j,k} =$ Expansión térmica lineal porcentual de la vaina evaluada a temperatura promedio de la vaina y pesado en volumen en $\Delta Z_{j,k}$.

$(V_{\text{Mod}})_{j,k} =$ Volumen del moderador en el incremento axial $\Delta Z_{j,k}$.

Análisis de resultados

Los cálculos obtenidos para el núcleo A16/21 son para inserciones de reactividad de \$1,5 a \$3 en 800 ms y son aplicados a dos modos extremos de operación establecidos en la rutina de trabajo del RP-10. Para el núcleo A16/21 se grafica la evolución del transitorio para el caso de refrigeración en convección forzada para \$3 y \$1,5 insertadas en 800 ms. De igual forma se ha calculado la evolución en el tiempo del transitorio en refrigeración por convección natural para \$1,5/800 ms. El comportamiento de este núcleo A16/21 para estos transitorios es comparado con los datos experimentales del reactor SPERT I para el caso de refrigeración forzada. En todos los casos se mantuvo el retardo de inicio de caída de barras en 200 ms. Este tiempo es bastante grande considerando el tiempo de 25 ms de los reactores SPERT I. Con este tiempo de retardo, en muchos transitorios, el núcleo ha evolucionado prácticamente sin protección. Es así que del análisis de estos resultados el caso más desfavorable es la inserción de \$3 en 800 ms. Esta excursión de potencia es la más violenta que puede ocurrir al RP-10 en la cual la temperatura de la vaina llega a temperatura de fusión a los 0,63 s después del segundo pico de potencia. Definitivamente aún cuando se tenga tres bombas en operación no se puede evitar la fusión de las placas si el sistema de protección banco de barras de seguridad no entra a mitigar este transitorio antes que ocurra este segundo pico de potencia. Aún cuando la señal está prefijada en la 12 MW, esta señal, enviada al sistema banco de barras de manera inmediata puesto que se supone que la electrónica asociada no tiene retardos, se da a los 0,4825 s pero el inicio efectivo de inserción de barras será después de 200 ms es decir a los 0,6825 s. Este tardío en introducir reactividad negativa de las barras de seguridad, permite que el sistema desarrolle un segundo pico de potencia a los 0,627 s con el cual el sistema libera una adicional energía considerable por el hecho de tener como pico ancho de potencia 207 MW llegando a tener como energía liberada total 18 MW-s. Este segundo pico lleva a que la temperatura de vaina llegue a la temperatura de fusión a los 0,63 s. Hasta este momento el reactor actúa totalmente desprotegido salvo por sus coeficientes de reactividad que son las que han permitido que la reactividad total del sistema se haga negativa. Esto se debe principalmente por la generación importante de burbujas en el primer pico de potencia. Las reactividades de realimentación por temperatura de moderador, de combustible y por efecto de dilatación se mantienen en valores de hasta -\$0,3. La realimentación por vacío induce una reactividad negativa de hasta -\$3,8 la cual es la causa principal de que la potencia no continúe incrementándose y la reactividad del sistema se hace negativa. Pero esta reactividad neta se vuelve positiva por dos razones, primero a que desafortunadamente la población de burbujas o porcentaje de vacío generado no se mantiene constante y por el contrario decrece, induciendo un efecto contrario a la seguridad del reactor por cuando es reactividad positiva ya que la reactividad de -\$3,8 que le quitó al sistema ahora tiende a devolvérselo, la segunda razón es que externamente todavía aún no se termina de insertar completamente la reactividad de \$3,0 causa del transitorio que al pico de potencia ya se insertó cerca de \$2. Otra de las causas principales de que el núcleo llegue a fundirse es que el núcleo no evacúa con

eficiencia el calor generado. Se puede ver que existe una mala refrigeración del núcleo y más aún cuando hasta el segundo pico de potencia el agua no ha sido capaz de retirar más de 6 MW-s de energía. Durante casi todo el transitorio el régimen de transferencia de calor que se mantiene en casi todas las cotas axiales es de ebullición en película o de transición. Las normas de seguridad reglamentan que para todos los casos accidentales no se debe superar el DNB. Este caso supera largamente esta limitación.

El análisis de los resultados para una inserción de \$ 1,5 en 800 ms manteniendo las mismas condiciones del caso anterior. Para ambos casos también se desarrollan dos picos de potencia pero con valores menores con respecto al caso de \$3/800 ms. Desde el punto de vista cálculo cinético el caso más desfavorable es con convección forzada por cuanto se consigue potencias de 285,6 y 50,4 MW para el primer y segundo pico de potencia respectivamente con 15 MW-s de energía generada al segundo pico de potencia, mientras que para el caso de convección natural se consiguen potencias 293,7 y 11,24 MW con 10,2 MW-s de energía generada. Pero de los cálculos termohidráulicos se aprecia que las temperaturas tienen un franco descenso para el caso de convección forzada llegándose en tiempos mayores de los 1,3 s a temperaturas de equilibrio de 20°C. Para el caso de convección natural la refrigeración es buena desarrollando el canal 1 regímenes de ebullición de bulbo de agua pero el sistema en su conjunto no retira el calor total generado con eficiencia es así que al segundo pico de potencia se ha retirado del núcleo solo 0,8Mw-s de energía. En el caso de convección forzada solo existe transferencia de energía por convección con el agua líquida. En ninguna cota axial se desarrolla regímenes de transferencia de calor más allá del DNB y el calor evacuado del sistema es bueno; tal es así que al segundo pico de potencia se retira del núcleo cerca de 7 MW-s y a los 1,35 s 15 MW-w. Esta buena evacuación del calor induce a que las temperaturas de los combustibles tengan un franco descenso habiéndose desarrollado una temperatura máxima de 430 y 390°C sólo justo después del pico de potencia. En el caso de convección natural se consigue temperaturas máximas de 500°C para la vaina y el combustible pero sin llegar a fundirse el núcleo. De los resultados analizados se puede concluir que el RP-10 en su configuración A16/21 sólo se puede admitir una inserción máxima de \$1,5 en 800 ms pero teniéndose en operación tres bombas. Es claro afirmar que de la forma como se ha corregido el tiempo de inicio de caída de barras a 80 ms el sistema de seguridad no permite que exista la evolución de un segundo pico y por tanto no permite mayor liberación de energía. Más aún las barras comenzarían a caer antes que termine el primer pico de potencia. La validez de estos resultados teóricos se comparan con los resultados experimentales de cuatro núcleos del reactor SPERT I. Se puede observar con claridad que el comportamiento del RP10 según el PARET guarda una buena concordancia con los datos experimentales para todos los casos en el rango de inserciones de \$1,5 a \$3 en 800 ms. Estas relaciones al pico de potencia son buenas por cuanto tienen similar carga de uranio los cuatro núcleos analizados salvo el SPERT-D12/25 que tiene 4200 grs. Esto se ve como las potencias son menores para todos los casos con respecto al RP-10. Con respecto a B12/64 y B24/32 estas tienen similar carga de uranio pero

las potencias del primer pico de potencia son mayores que las del RP-10. Esto es consecuencia de que el bajo valor de los coeficientes de reactividad por vacío, con respecto al del RP10-A16/21, permiten una mayor evolución de potencia en estos núcleos.

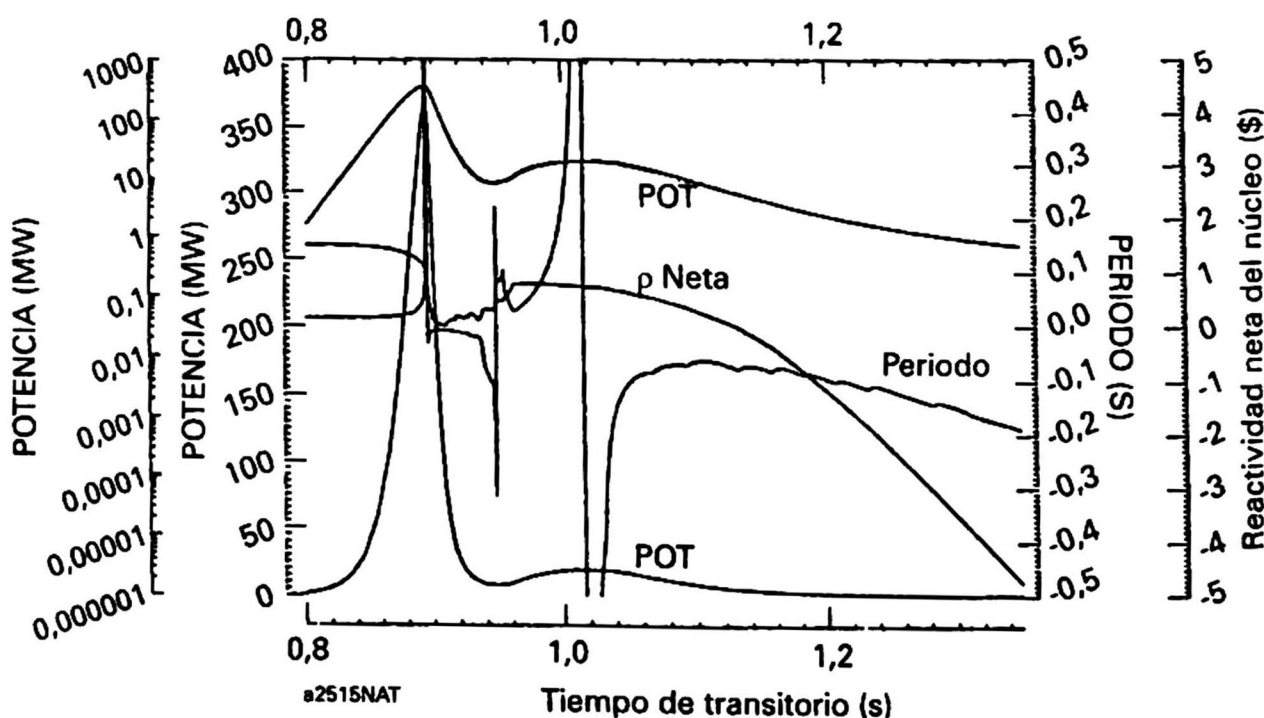
Para el análisis de este núcleo 16 ó A16/25, el caso de convección natural, es el caso más desfavorable con respecto al de convección forzada. Para el caso de convección forzada se tuvo problemas en la subrutina tabla 1 para el cálculo de las propiedades del líquido subenfriado de forma tal que no se podía continuar con el cálculo de la evolución del transitorio. Por analogía con el núcleo 14 se reporta lo que puede suceder al núcleo 16 en convección forzada. El estudio de seguridad de éste núcleo queda cubierto en razón de que si para el núcleo 14 el caso más desfavorable es trabajar en modo de convección natural, para el núcleo 16 es también el caso más desfavorable. En éste caso se desarrolla una temperatura máxima de 380°C pero es sensiblemente menor de los 590°C que se consigue en las mismas condiciones accidentales con el núcleo 14. Si bien el núcleo 16 tiene un inventario de Uranio 235 mayor, 7105 grs., la energía liberada total acumulada se reparte de forma más homogénea en todo el núcleo. Esto se puede deducir cuando se tiene como primer canal cuatro elementos y como factor pico solo 2,32; lo cual favorece en gran medida que el calor no se concentre en un solo elemento como sucede en el núcleo 14. En todos los casos las energías liberadas al pico de potencia son mayores en el caso de convección forzada con respecto al de convección natural pero el sistema es muy eficiente para retirar el calor del núcleo cuando se tiene operando las tres bombas. Para estos transitorios también se ha supuesto tener el peor caso con un retardo de 200 ms. Teniendo el sistema de caída de barras con un retardo de 80 ms se puede observar que las barras deben empezar a caer antes de que termine de evolucionar el primer pico de potencia. Si ha esto se incluye la disposición de que necesariamente se debe tener las tres bombas operando para cualquier condición el reactor evolucionará favorablemente sin tener de consecuencias una posible fusión del núcleo.

El RP-10 al tener el modo de operación de trabajar a bajas potencias en convección natural, se postula para esta condición que ocurre el accidente de inserción de $1,5/800$ ms. Se puede observar los resultados de aplicar esta inserción en el gráfico I con un retardo de inicio de caída de barras de 80 ms. Se observa que las temperaturas tienen un franco descenso hacia las temperaturas de saturación del canal 1. La máxima temperatura que se desarrolla son 354 y 350°C para el combustible y la vaina respectivamente. Este pico de temperatura se desarrolla después de 0,9 s y durante el desarrollo de este pico en los dos canales se desarrolla el régimen de transición que es superado desarrollándose posteriormente en el canal 1 ebullición de bulbo de agua y el canal 2 ebullición nucleada. Este tipo de transferencia de calor permite que haya una buena refrigeración de ambos canales y en consecuencia las temperaturas descenden. Con respecto al segundo pico de potencia, éste se desarrolla pero de manera mucho menos pronunciada con respecto al segundo pico que se consigue con un

retardo de 200 ms. Este efecto se debe principalmente a que la reactividad positiva del sistema comienza a ser compensada con la reactividad que introduce el banco de barras de seguridad desde los 0,91 s. Aún cuando a partir de este momento las burbujas tienden a disminuir introduciendo reactividad positiva las barras de seguridad no llega a compensarlas de manera efectiva en razón de que los primeros cms. de las barras introducen muy poca reactividad. Esta reactividad positiva inducida por la desaparición de las burbujas y el descenso de temperatura del combustible es compensada por las barras de seguridad pero se logra desarrollar un segundo pico que es anulada por la continua introducción de las barras de seguridad.

**Excursión de potencia por inserción de reactividad
RP10-A16/25 (Núcleo 16)**

$\beta = 1,5/800$ ms, $\tau = 200$ ms, 1C4E, CNat, $P_0 = 5$ watts, Scram: 12 MW, $1/T_0 = 67,9$ (1/s)



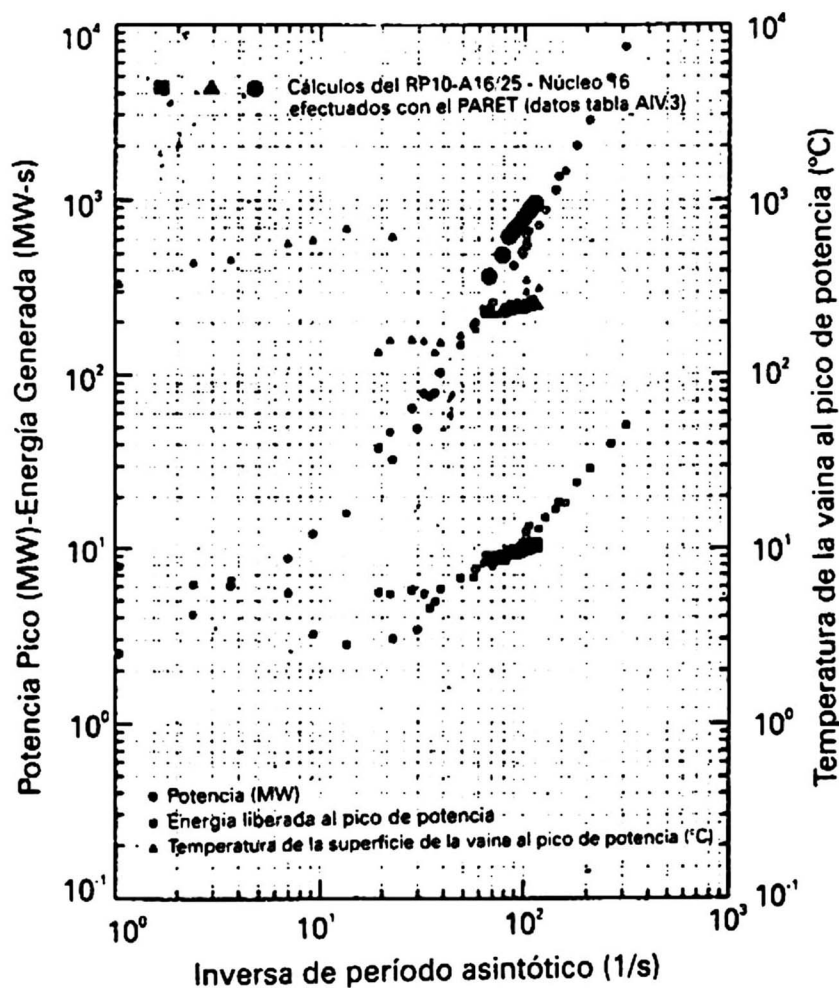
Se concluye también que el caso más desfavorable es cuando se tiene un beta efectivo mayor para el cual para las mismas condiciones se tiene una potencia pico de 402 MW. Las energías liberadas al pico de potencia son similares. Una respuesta a este efecto es que si bien la potencia pico es mayor al ancho del pico de potencia se hace más angosta. Como las temperaturas son función de la energía liberada, estas se mantienen casi en los mismos valores al menos para la vaina. El combustible si experimenta una apreciable aumento de temperatura. Otro aspecto en contra de la seguridad es que el pico de potencia se manifiesta en tiempos más cortos, es así que

mientras se tiene el pico de potencia a los 0,89 s para beta de 0,0072; este pico de potencia aparece a los 0,66 cuando se corre el caso con un beta de 0,0078. En este último caso las reactividades que tendrá que compensar las barras de seguridad son las de vacío, la de descenso de temperatura de combustible y además la reactividad externa que termina de ser insertada sólo después de los 800 ms.

Conclusiones

Los núcleos RP10-A16/21 y A16/25, simulados con el PARET tienen muy buena concordancia con los resultados experimentales de los reactores SPERT-I (fig. 2). El núcleo SPERT I de bajo enriquecimiento es el que representa de forma más precisa la conducta del RP10 en su última configuración A16/25 ó núcleo 16.

Cálculos del RP10 y datos del SPERT-I: 4,02Wt% U-235



Se concluye que trabajar con reactividades mayores a $\$1,5/800$ ms traería de consecuencia dificultades en extraer el calor generado en la excursión y por lo tanto no es posible controlar el incremento de la temperatura del combustible y de la vaina.

El modo de operación más desfavorable del RP10 es trabajar con refrigeración por convección natural, al menos para este tipo de accidentes de excursiones de potencia por inserciones de reactividad. Para este modo de operación la configuración 16 puede soportar en refrigeración por convección natural una tasa de inserción de \$1,5 en 800 ms como máximo. Se observa que esta reactividad insertada en tiempos menores conlleva a que el pico de potencia suceda antes y con picos de potencia más elevados.

Es importante efectuar en forma rutinaria una evaluación experimental de los tiempos de retardo de inicio de caída de las barras de seguridad, por cuanto se demuestra la gran importancia que tiene éste en el desarrollo y el control de la reactividad total del sistema y por lo tanto del control del transitorio hacia un estado seguro del núcleo.

De igual forma es importante efectuar un control frecuente del peso de las barras de seguridad por cuanto se observa que la reactividad total de las barras de seguridad se hace menor para núcleos más grandes tal como el 16 con respecto al núcleo 14. Si bien se tiene un buen margen de reactividad disponible para mitigar o compensar la reactividad causa del transitorio es fundamental conocer el peso de reactividad de los primeros tramos de las barras de seguridad que empiezan a ser introducidas al núcleo.

Es importante que los coeficientes de realimentación sean calculados experimentalmente de manera frecuente y con rutinas experimentales más precisas en virtud de que son los que actúan de manera inmediata para evitar que la potencia se incremente sin control. Los coeficientes de vacío y de temperatura de moderador medidos experimentalmente para el núcleo 16 son menores con respecto al núcleo 14. Este descenso es perjudicial para la seguridad del reactor por cuanto se induce menor reactividad al pico de potencia. Se optimizó los sistemas de seguridad del reactor para llevar al núcleo a una parada segura. Se consiguió [5] disminuir la corriente de los electroimanes para permitir una más rápida desenergización y una apertura de los relays con menor retardo de manera que se inicie la caída de las barras de seguridad.

REFERENCIAS

- [1] C. F. Obenchain, *PARET. A program for the Analysis of Reactor Transient*, IDO-17282 (January 1969).
- [2] W. L. Woodruff, *PARET/ANL. Argonne Code Center, Master Copy Abstract 555*.
- [3] R. S. Smith, "PARET code revisions", Argonne National Laboratory (December 1987).
- [4] G. S. Lázaro. "Implementación del programa PARET a una PC". Informe Técnico 013/DCAS/DGI, Instituto Peruano de Energía Nuclear (Noviembre 1993).
- [5] A. Ramón. "Medición del tiempo de retardo de inicio de caída de barras de los sistemas de seguridad del RP-10". Memo s/n, Instituto Peruano de Energía Nuclear (Julio 1993).

- [6] W. L. Woodruff. *"A kinetics and Thermal-Hydraulics capability for the analysis of research reactors"*, Nuclear Technology Vol 64 (February 1984).
- [7] A.P. Wing. *"Transient tests of the fully enriched, Aluminum Plate-type, B cores in the SPERT I Reactor Data summary report"*, IDO-16964, Idaho National Engineering Laboratory (June 1964).
- [8] M.R. Zeissler, *"Non-destructive and destructive transient tests of the SPERT I-D, fully enriched, Aluminum Plate-type core: Data summary report"*, IDO-16886, Idaho National Engineering Laboratory (Nov. 1963).

Síntesis y estudio de complejos de 3-(2-Piridil)-5,6 -Difenil-1, 2, 4- Triazina con hierro (II) y cobre (I), su determinación espectrofotométrica y aplicación al análisis de agua pre-tratada de uso industrial

Patricia Rosario López Pino (*)

RESUMEN

La síntesis del reactivo cromogénico 3-(2-piridil)-5,6-difenil-1,2,4-triazina permite revisar la formación de los complejos de Fe (II) y Cu (I) con el mencionado ligando (PDT), ampliar la caracterización de los complejos y verificar su determinación espectrofotométrica simultánea en soluciones estándar de hierro y cobre. El método se aplica finalmente a la determinación de ambos iones en agua pre-tratada para sistemas de enfriamiento industrial.

Introducción

F. Blau [3] reportó la preparación de 2,2'-bipiridilo en 1888 y de 1,10 - fenantrolina en 1898, descubrió también sus reacciones coloreadas con Fe (II). A partir de 1930 recién se les usa como reactivos cromogénicos para hierro y como indicadores. En la segunda mitad de la década de 1940 se crea un programa colectivo de estudio sistemático entre las universidades de Illinois y Temple, con la

(*) Facultad de Ciencias, Universidad Nacional de Ingeniería, Casilla postal 31-139, Lima - Perú.

finalidad de sintetizar este tipo de compuestos. Consecuencia de este programa es la batofenantrolina, TPTZ (2, 4, 6-trispiridil-s-triazina) y PDT entre otros.

Todos estos compuestos se caracterizan por la presencia del grupo α, α' -diimina ($\text{-N}=\overset{\text{I}}{\text{C}}-\overset{\text{I}}{\text{C}}=\text{N-}$) cuyos nitrógenos se unen al ión metálico, formando un anillo de cinco miembros. Los complejos así formados son del tipo amina. Los átomos de hidrógeno no son reemplazados y la extracción en un solvente inmiscible requiere de la participación de un anión que forme una asociación o par iónico.

Parte experimental

Síntesis de PDT [1]

Reactivos:

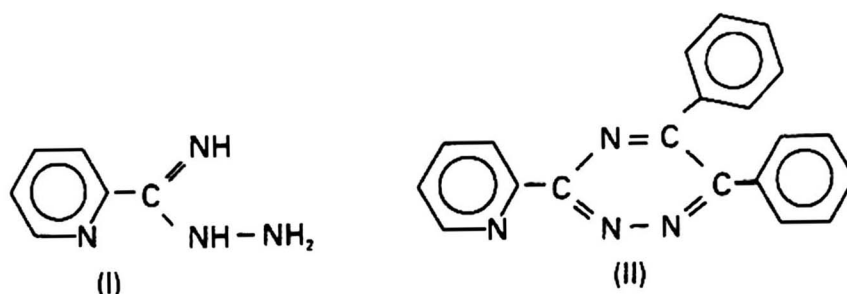
Hidracina al 64% ($\text{N}_2\text{H}_4 \cdot \text{H}_2\text{O}$); 2-cianopiridina; etanol absoluto; éter dietílico; tolueno; bencilo.

Preparación de Hidracidina:

Se agita por dos horas y a temperatura ambiente, una mezcla de 20 ml de hidracina con 0,05 mol de cianopiridina en 9 ml de etanol. Los cristales formados de 2-piridilhidracidina (I) se extraen en éter y se secan sobre sulfato de sodio anhidro. Después de destilar el éter, el residuo se recrystaliza en tolueno.

Preparación de la triazina asimétrica trisustituída PDT (II):

Se mezcla 0,05 mol de (I) con 0,05 mol de bencilo en 25 ml de etanol. Se deja reposar por 48 horas o hasta que se evapore el etanol. Los cristales aciculares amarillos se recrystalizan en etanol hasta obtener punto de fusión constante.



Caracterización de PDT:

Se determina el punto de fusión por el método gráfico de Shriner y Fuson, obteniéndose el valor de 190,7 C (valor tabulado: 189-190C)¹.

Registro del espectro IR (equipo Perkin Elmer, modelo 882-PUC).

Síntesis y caracterización de complejos de Fe (II) y Cu (I) con PDT [2]

Reactivos:

PDT 0,01 M en etanol (2 gts HCl 12M/25 ml de solución); soluciones patrón de hierro y cobre electrolíticos; cloruro de hidroxilamonio 10%*; solución buffer* (pH 4,5); perclorato de sodio* 50%; ácido ascórbico 0,5% en etanol (preparar antes de usar); cianuro de sodio; alcohol isoamílico.

(*) *Extraer con PDT en alcohol isoamílico (previo ajuste con reductor, buffer o perclorato).*

Equipo:

Espectrofotómetro Perkin-Elmer Lambda 3A UV/VIS (IBEMETSA)

Síntesis:

En fioles de 10 ml se prepara individualmente los complejos con 10,69 µg Fe y 43,3 µg Cu. Se añade en ambos casos: 2 ml NH₂OH . HCl 10%, 2 ml PDT 0,01 M, 2 ml NaClO₄ 50%; se extraen con 6 y 2 ml de alcohol isoamílico, se enrasa con etanol.

Resultados

Tabla I. Características de los complejos de Fe (II) y Cu (I) con PDT

	Fe (PDT) _n ²⁺	Cu (PDT) _m ⁺
Color	rojo	marrón
Curvas espectrales (transferencia de carga)	Fe ²⁺ (3d ⁶) t _{2g} → π*	Cu ⁺ (3d ¹⁰) e → π*
λ max.	556 nm	490 nm
Solubilidad como percloratos	- en agua insoluble - en alcohol isoamílico soluble	insoluble soluble
Estabilidad en fase orgánica, (sin añadir un 2º reductor).	estable	inestable (*)
Elección de reductores y rango óptimo de pH	pH A _{MAX} NH ₂ OH.HCl (1) 1 - 3 0,5 Ac. ascórbico (2) 1,5 - 6 0,6 (1) y (2) (Fig.1) 1,5 - 6 0,6	pH A _{MAX} 5 - 7 0,7 1,5 - 7,5 0,9 3 - 6 1,0
Estequiometría (Método de variaciones continuas)	n = 3	m = 2

(*) *Precisa la adición de un agente reductor después de la extracción en alcohol isoamílico.*

Fig. 1-a Curvas espectrales de Fe (II) a diferentes pH

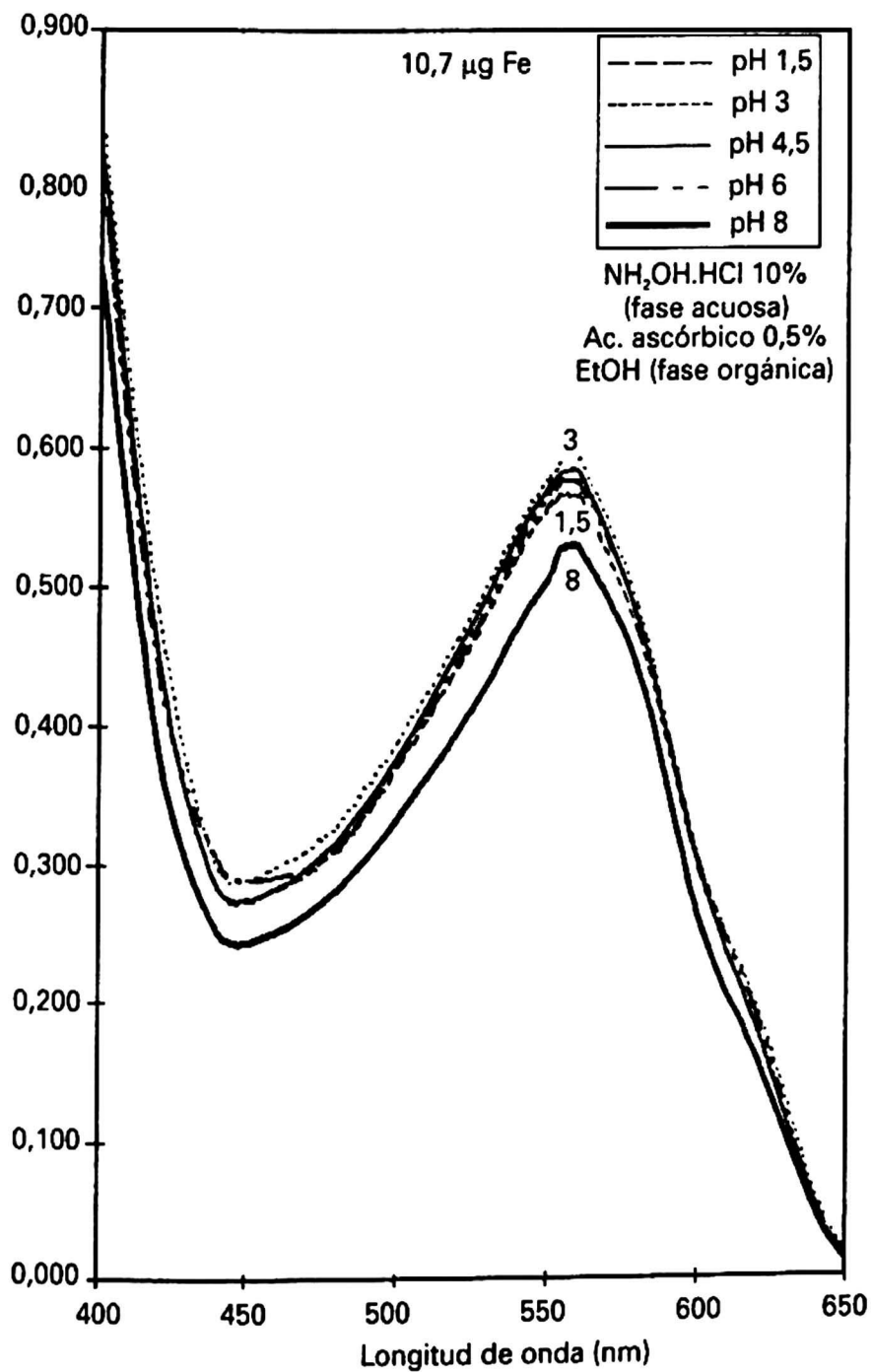
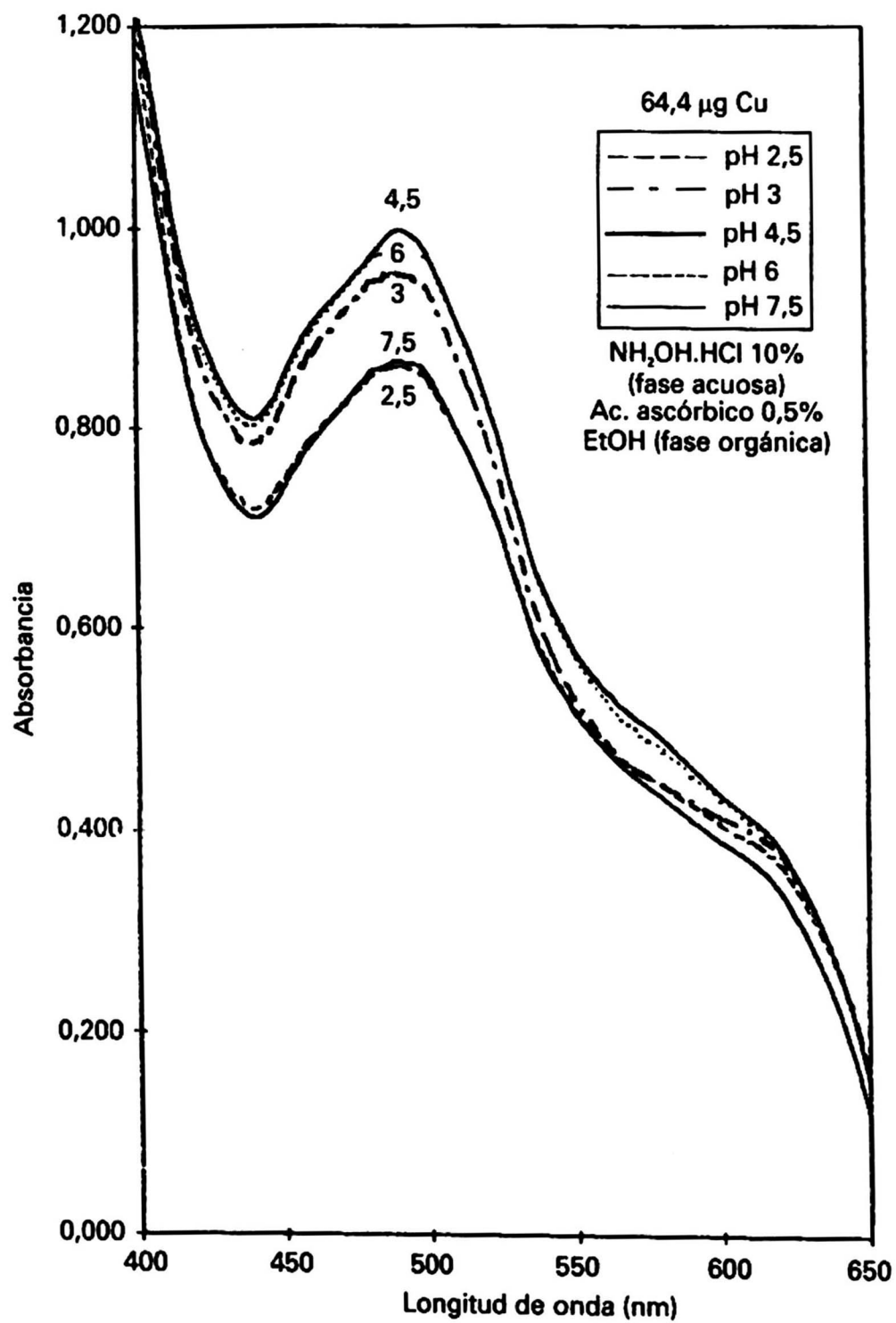
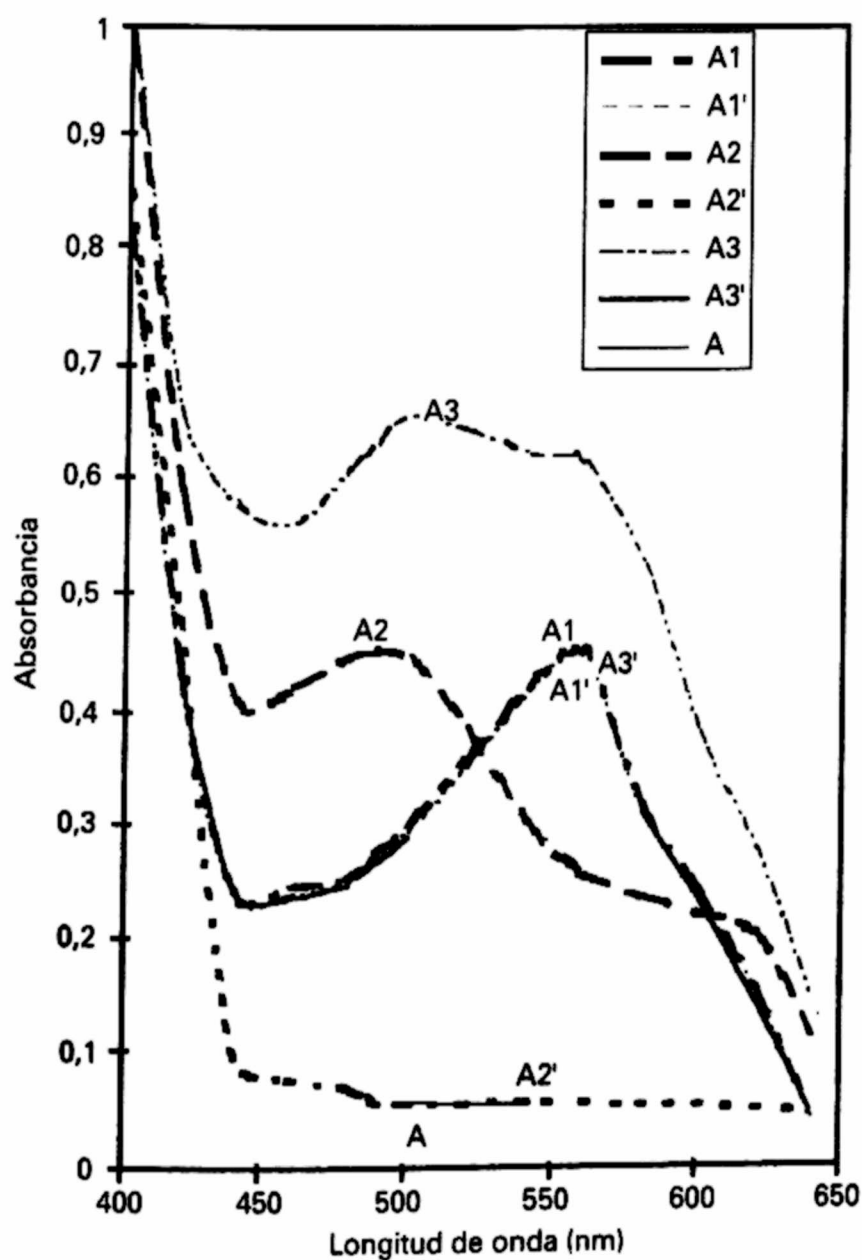


Fig. 1-b Curvas espectrales de Cu (I) a diferentes pH

Efecto de la adición de cianuro de sodio en los complejos de Fe (II) y Cu (I) con PDT

A los complejos formados, extraídos en alcohol isoamílico y estabilizados con ácido ascórbico, se añade cianuro de sodio seco y se observa la destrucción del complejo Cu (I) - PDT. (Fig. 2).

Fig. 2 Curvas espectrales de los complejos de Fe y Cu con PDT antes y después de añadir NaCN



- A1 : Fe-PDT (8,6 μg Fe);
 A2 : Cu-PDT (25,2 μg Cu);
 A3 : Fe y Cu con PDT
 A : BK + NaCN
- A_i' : después de añadir NaCN
 (i = 1, 2, 3)

Determinación analítica de estándares

Procedimiento recomendado:

En un embudo de separación de 100 ml añadir, una alícuota de hierro y/o cobre disueltos, 5 ml de amortiguador (pH 4,5), 2 ml cloruro de hidroxilamonio 10%, 2 ml PDT 0,01 M, 2 ml NaClO_4 50%. Chequear el pH ($4,5 \pm 1$) con papel indicador, extraer con 6 y 2 ml de alcohol isoamílico, enrasar con ácido ascórbico 0,5% en etanol. Medir absorbancia a 490 nm. Añadir NaCN (5 a 10 mg) agitar fuertemente, medir nuevamente la absorbancia a 490 nm y luego a 556 nm.

Efecto de tiempo

La formación de los complejos es instantánea y las absorbancias deben registrarse a más tardar a una hora de extraídos los complejos, debido principalmente al complejo de Cu (I), catión lábil (d^{10}).

Reproducibilidad

Se prepara tres curvas estándar, en días diferentes y se obtiene valores de absorbancia que difieren entre sí en $\pm 0,002$ unidades de absorbancia, en rangos de 0,2 a 1,6 ppm de hierro y de 0,4 a 3,2 ppm de cobre.

Ley de Beer

Intervalo óptimo de Ringbom - Absortividad.

Se comprueba la ley de BEER para ambos complejos, según el punto de inflexión en el gráfico de Ringbom (37% absorbancia).

El intervalo óptimo de concentración es: (0,35 - 1,6) ppm Fe y (0,8 - 7,0) ppm Cu.

El porcentaje de error relativo de concentración en el rango establecido es de 1,02% (Fe-PDT) y de 0,94% (Cu-PDT), considerando un error fotométrico de 0,3% T.

La absortividad molar media es de $24127 \text{ l.mol}^{-1} \cdot \text{cm}^{-1}$ para el hierro y de $7887 \text{ l.mol}^{-1} \cdot \text{cm}^{-1}$ para el cobre.

Sensibilidad de SANDELL:

0.0023 y 0,008 $\mu\text{g.cm}^{-2}$ para Fe y Cu respectivamente.

Aplicación a la determinación de hierro y cobre en agua pre-tratada de uso industrial

El agua pre-tratada de dureza 41 ppm CaCO_3 proviene de la mezcla de agua blanda (dureza cero) con agua dura. El procedimiento previo al análisis de las muestras de agua es: se colecta 100 ml de agua pre-tratada y se añade 4 ml HCl 12M (Norma ASTM D 3370-76); se evapora sin hervir con 3 ml HNO_3 15M, se lleva a sequedad y se enfría. Se diluye con 0,5 ml HCl 12M y se lleva a volumen de 10 ml.

Los resultados de la determinación espectrofotométrica simultánea son: $(0,093 \pm 0,009)$ ppm Fe y $(0,249 \pm 0,022)$ ppm Cu.

Como patrón de comparación se analiza otras alícuotas de la misma muestra de agua pre-tratada (acidificada y concentrada de la misma manera) por espectrofotometría de absorción atómica, dando como resultados: 0,096 ppm Fe y 0,292 ppm Cu.

Interferencias con otros iones

El análisis semi-cuantitativo de los principales interferentes arroja los siguientes resultados: Ni (II) < 3 ppm, Cr (III) < 100 ppm, NO_2^- < 100 ppm, $\text{C}_2\text{O}_4^{2-}$ < 50 ppm, $\text{S}_2\text{O}_3^{2-}$ < 20 ppm, MoO_4^{2-} < 25 ppm.

Ninguno de los cuales sobrepasa el límite determinado, según el cual se les tendría que aislar o enmascarar.

El Co (II), principal inteferente de Cu (I) no fue analizado, dada su ausencia en aguas de Lima, según Sedapal.

Costos

Síntesis de PDT:

1 mol 2 - cianopiridina (~ 100 g)	\$ 15,00
400 mol hidracina	40,00
1 mol bencilo (~ 210 g)	15,00
2ℓ etanol	20,00
1,5 ℓ éter	25,00
1 ℓ tolueno	20,00
	<hr/>
	\$ 135,00 (total)

Se obtiene: $1 \text{ mol PDT} \times 310 \text{ g/mol} \times 0,12 = 35 \text{ g PDT}$.
(Considerando el 12% de rendimiento experimental)

Determinación de Fe y Cu con PDT:

10 g PDT se usa en 1600 determinaciones individuales	
10 g PDT	\$ 40,00
3,2 l etanol	30,00
13 l alc. isoamílico	400,00
	<hr/>
	\$ 470,00 (total)

\$ 0,30 por determinación simultánea de una muestra de Fe y Cu.

Determinación individual de Fe con biperidilo:

3,2 g bipi en 1600 determinaciones individuales.	
3,2 g bipi	\$ 11,00

Determinación individual de Cu con DTC (Dietilditiocarbamato de sodio):

En 1600 determinaciones individuales,	
4 g DTC	\$ 9,00
2K EDTA (sal Na)	200,00
16 l CCl ₄	100,00
	<hr/>
	320,00 (total Fe y Cu)

\$ 0,20 por determinación de Fe y Cu (independiente una de otra).

Disposición final de reactivos tóxicos

Todos los reactivos utilizados en el presente trabajo, excepto el alcohol isoamílico, se encuentran inventariados en la Agencia de Protección del Ambiente (EPA) de los E.U.A. (Acta de Control de sustancias tóxicas).

Los residuos de hidracina, nitrilo, bencilo y PDT deben quemarse con un solvente combustible en un incinerador provisto de quemador auxiliar y depurador de gases.

Los residuos de etanol, éter, benceno o tolueno, deben quemarse en pequeñas cantidades y con sumo cuidado (volumen de la sustancia no exceda el 50% de la capacidad del recipiente que lo contiene al ser incinerado).

Uso de lentes, guantes y ropa de protección y en todos los casos usar el mismo tipo de incinerador químico descrito.

Conclusiones

- 1) La síntesis del reactivo PDT fue ligeramente modificada, debido a la imposibilidad de conseguir hidracina al 95% y al uso de solventes como tolueno en lugar de benceno; reportándose un producto con las mismas características que el de la síntesis original, aunque con aparente menor rendimiento.
- 2) La aplicación analítica del método espectrofotométrico en el agua de uso industrial, permite afirmar que el método tiene una respuesta bastante aceptable, con una desviación menor al 10% en los resultados obtenidos.
- 3) El método de determinación simultánea de Fe y Cu resulta ventajoso por:
Menor tiempo de ejecución.
Menor contaminación.
Menor costo en la síntesis de PDT y en la adquisición de equipos instrumentales necesarios.

REFERENCIAS

1. F. H. Case, *J. Org. Chem.* 30, 931 (1965).
2. A. Schilt, P. Taylor, *Anal. Chem.* Vol. 42, N° 2, Feb. 1970, 220-224.
3. A. Schilt, W. Hoyle, *Anal. Chem.* Vol. 39, N° 1, Jan. 1967, 114-117.

Agradecimientos

*A CONCYTEC (Oficina de apoyo al investigador).
A Gina Chiarella y Mario Ceroni.
Al Instituto de Investigación de la Facultad de Ciencias.
A la fábrica Iberandina de Metales S.A. (IBEMETSA).*

Estudio de las propiedades estructurales ópticas y eléctricas de $\text{Sn O}_x\text{: F}$ obtenido de soluciones pirolíticas con alto contenido de flúor

Enrique Montoya Quiroz (*)

RESUMEN

En el presente trabajo, se han obtenido películas de $\text{Sn O}_x\text{: F}$ mediante el método de rociado pirolítico. Los espectros de rayos x muestran una estructura casiterite Sn O_2 con un crecimiento preferencial en la dirección (200).

Las micrografías (SEM) muestran que la superficie de las películas no dopadas poseen partículas cristalinas cuyo tamaño es de $0,5 \mu\text{m}$. Al aumentar la concentración de NH_4F , se evidencia un sistemático decrecimiento en la densidad de partículas superficiales; la transmitancia se incrementa con la concentración llegando a 92% para 20% F/Sn.

(*) Facultad de Ciencias - Universidad Nacional de Ingeniería, Casilla Postal 31-139, Lima - Perú.

La resistividad decrece a $0,42 \times 10^{-3} \Omega\text{cm}$ para rangos de dopaje de 1-8%. Los estudios de difracción de rayos X indican que para una concentración de 30% F/Sn, los átomos de flúor ocupan mayormente posiciones sustitucionales de oxígeno. Pero no se descartan las posiciones intersticiales aunque en menor proporción.

Introducción

El SnO_2 es un material muy interesante debido a sus múltiples aplicaciones tales como: sensor de gas, espejos de calor y conversión fotovoltaica de la energía solar. La conductividad está dada por deficiencias de oxígeno las que producen niveles de ionización ubicados muy cerca del fondo de la banda de conducción. Un factor que afecta esta conductividad del SnO_2 son los niveles de impureza proporcionados en este caso por dopaje con flúor. Las películas de óxido de estaño se preparan con muchos métodos, algunos de los cuales son: pulverización reactiva, evaporación y rociado pirolítico de compuestos estánnicos, etc. siendo esta última, una buena alternativa para aplicaciones en áreas grandes.

Los estudios reportados en la literatura acerca del SnO_2 : F [1] enfocan principalmente sus propiedades ópticas y eléctricas. Los cambios estructurales dados por la incorporación de flúor en las películas policristalinas no han sido extensamente discutidas. Fantini [2] por ejemplo, ha reportado estudios de la influencia del flúor en las características estructurales, pero sus conclusiones se han basado principalmente en los datos de difracción de rayos X.

Esto ha motivado la realización del presente trabajo, en el cual se proponen los siguientes objetivos.

- Estudiar la influencia del dopaje con altas concentraciones de flúor en las características morfológicas, estructurales, ópticas y eléctricas.
- Proveer de más elementos experimentales que permitan entender el mecanismo por el cual el flúor altera las propiedades físicas del SnO_2 .
- Realizar los primeros estudios que permitan mejorar la capacidad de respuesta del SnO_2 como sensor de gas a partir de modificaciones morfológicas de la película; particularmente incrementando la relación superficie volumen.

Detalles experimentales

Las películas de SnO_2 : F han sido producidas por un sistema spray, una temperatura de sustrato $T = 340^\circ\text{C}$ y un tiempo de depósito de 15 minutos. La

solución consiste de una sal de Sn Cl_4 disuelta en alcohol a una concentración de 0,2 M. Se añadieron además diferentes concentraciones de NH_4F para conseguir películas con diferente concentración de dopante.

Las medidas de transmitancia óptica se hicieron en un Espectrómetro OPTOMETRIC RS-375 en el rango de 0,32 a 0,85 μm . De estos datos se ha calculado también el espesor de las películas. La resistencia superficial se ha medido siguiendo los procedimientos convencionales. Los datos de difracción de rayos X fueron tomados en un Difractómetro Zeiss Jena a 30 KeV y 20 mA, usando la radiación $\text{K}\alpha$ ($\lambda = 1,542 \text{ \AA}$) del cobre.

Las micrografías de Microscopía Electrónica de Barrido (SEM) fueron tomadas en un microscopio electrónico JEOL 5200.

El contenido de flúor en las películas se ha medido por el método de Reaction Nuclear Resonante (RNR). En este método, un haz de protones de 700 KeV de un acelerador de Van de Graaf excita una reacción nuclear $^{19}\text{F} (p, \alpha \gamma) ^{16}\text{O}$ de 340 KeV. El rayo gamma de 6,14 MeV emitido por la reacción es detectado en un detector de centelleo NaI 76 x 76 mm (T ℓ). La curva de reacción excitación ha sido ajustada usando una expresión conocida [4], con el fin de encontrar el perfil de la concentración de flúor en la muestra.

Resultados

La tabla 1 muestra la concentración de flúor en solución, la resistividad eléctrica y el espesor de las muestras de óxido de estaño obtenidas.

Tabla 1

% (F/Sn)	Espesor	$\rho \cdot 10^{-3} \Omega \text{ cm}$
0	368	3,31
1,2	363	1,31
2,3	482	1,03
3,4	410	0,77
4,5	303	0,53
5,7	264	0,47
6,8	351	0,42
8,0	427	0,76
15	195	1,20
20	196	2,58

INCORPORACIÓN DEL DOPANTE

La fig. 1 muestra la profundidad alcanzada por la distribución de flúor obtenida por RNR, respecto a la concentración de flúor en la solución. A bajas concentraciones la distribución de flúor es profunda; de acuerdo a como la concentración de flúor en la solución se incrementa, la distribución de flúor respecto a la profundidad decrece y su valor mínimo se encuentra alrededor de 20% F/Sn.

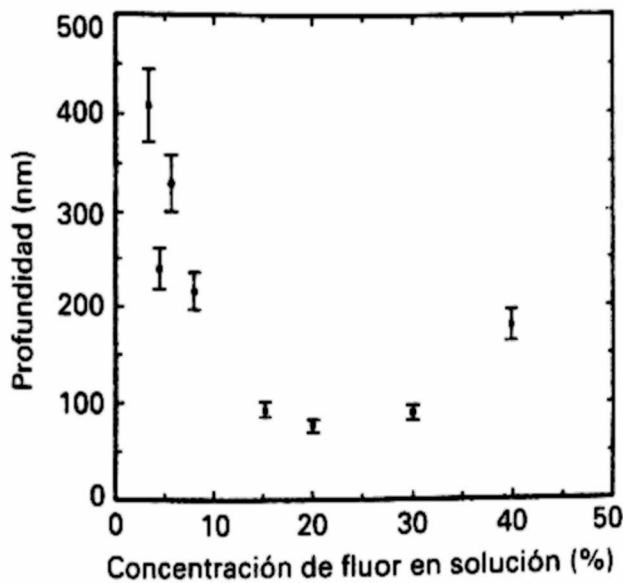


Fig. 1
Profundidad de la distribución de flúor en función al porcentaje de flúor en la solución.

En la figura 2 se muestra la distribución de flúor en la película como función de la concentración en la solución. Se puede deducir una dependencia lineal entre ambos. La línea recta corresponde al ajuste de mínimos cuadrados con un factor de correlación de 0,902.

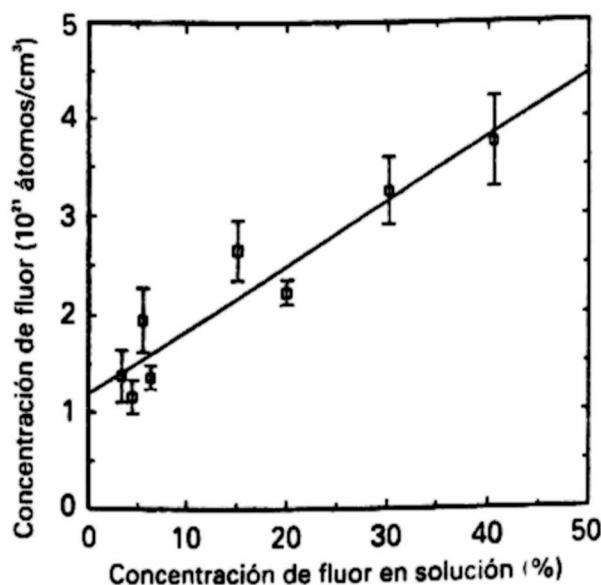


Fig. 2
Concentración de flúor en la película en función de la concentración de flúor en la solución.

PROPIEDADES ÓPTICAS Y CARACTERÍSTICAS MORFOLÓGICAS

La figura 3 muestra las curvas de transmitancia espectral que corresponde a diferentes niveles de dopaje. Se puede observar que la transmitancia de estas películas se incrementa a medida que la concentración de flúor en la solución se incrementa. La transmitancia óptica para películas no dopadas es de 75% (650 nm, curva figura 1) y se incrementa hasta 92% (curva d figura 1) cuando la muestra es dopada con 20% F/Sn (peso).

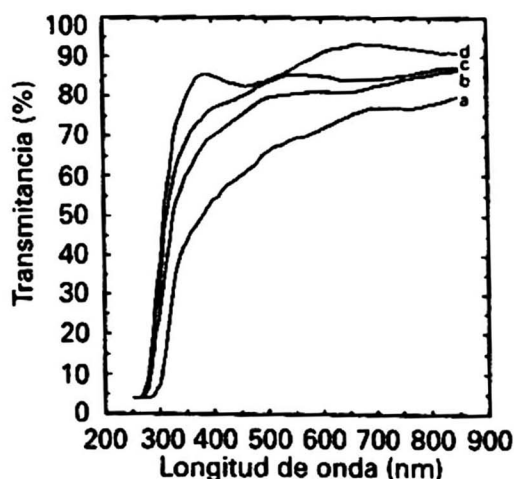


Fig. 3
Transmitancia espectral para
películas de $\text{SnO}_2 : \text{F}$
(a) 0%, (b) 8%,
(c) 15%, (d) 20% F/Sn

PROPIEDADES ELÉCTRICAS

La figura 4 muestra la resistividad de las películas de SnO_2 para diferentes contenidos de flúor. Las muestras sin dopaje presentan una resistividad de $3,4 \times 10^{-3} \Omega \text{ cm}$. Para bajos niveles de dopaje (1% al 8%) la resistividad llega a valores de $0,42 \times 10^{-3} \Omega \text{ cm}$. Para altos niveles de dopaje (mayores que 8%) la resistividad muestra un progresivo incremento: muestras con 15% y 20% F/Sn tienen una resistividad de $1,20 \times 10^{-3}$ y $2,58 \times 10^{-3} \Omega \text{ cm}$, respectivamente.

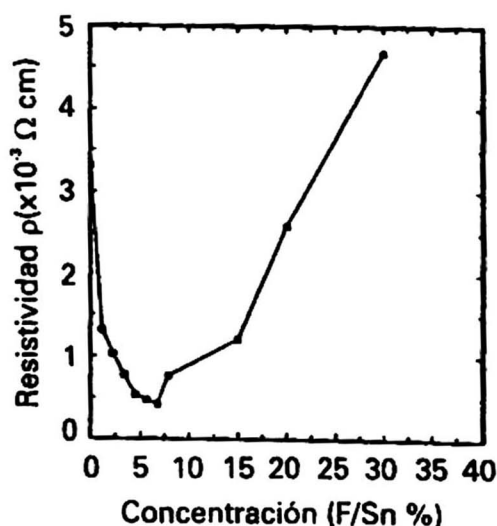


Fig. 4
Resistividad en función de la con-
centración de flúor en la solución
para películas de $\text{SnO}_2 : \text{F}$.

DETALLES ESTRUCTURALES

La figura 5 muestra los espectros de difracción de rayos X para películas de óxido de estaño, las que han sido obtenidas con diferentes niveles de dopaje. La figura 5 curvas a, b y c, corresponden a muestras con 0%, 4,5% y 8% F/Sn, respectivamente; se puede observar que la intensidad de los picos (211), (200) y (110) decrece sistemáticamente y el pico (101) permanece sin cambio apreciable.

Este comportamiento varía cuando el nivel de dopante se incrementa; la figura 5d muestra el difractograma para una muestra con 30% F/Sn. Aquí el pico (101) decrece, y los picos (200) y (110) se incrementan. A este nivel, se observa también la aparición de un nuevo pico a un ángulo $2\theta \approx 31,5^\circ$.

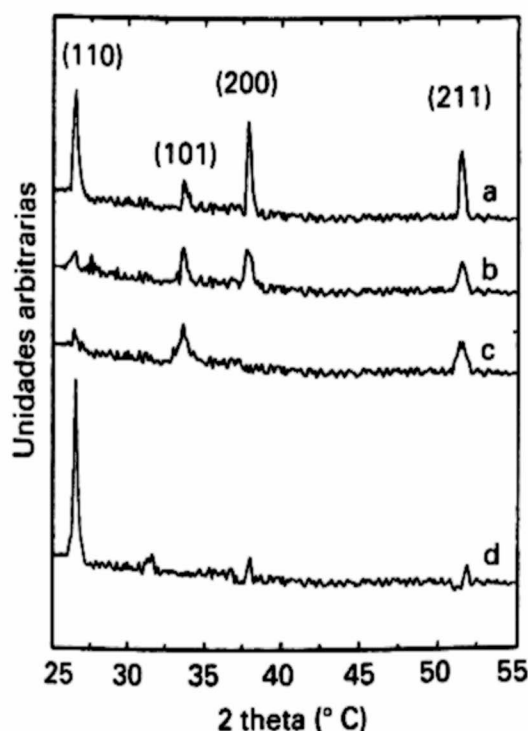


Fig. 5
Espectros de difracción de rayos X de $\text{SnO}_2 : \text{F}$
(a) muestra no dopada F/Sn 0%
(b) muestra dopada F/Sn 4,5%
(c) muestra dopada F/Sn 8%
(d) muestra dopada F/Sn 30%

Discusión

Al incrementarse la concentración del dopante (0-8%), la densidad de cristalitos en la película disminuye. Para 8% F/Sn, la superficie de la película es muy lisa, este decrecimiento en la rugosidad produce una disminución en la dispersión de la luz y en consecuencia, la transmitancia óptica se incrementa (fig. 1). Para 20% F se observa la presencia de algunas picaduras.

Para bajos niveles de dopante (menos que 8%) la resistividad eléctrica disminuye hasta en un 90% respecto a las películas no dopadas; esto se atribuye a un incremento

en el número de portadores libres. Sin embargo, cuando el dopante es mayor (más que 10%) la resistividad se incrementa (más de diez veces para 30% F/Sn); esto puede ser debido a la formación de picaduras en la película.

Del análisis de los espectros de difracción de rayos X, se ha encontrado para películas no dopadas una estructura casiterita con un crecimiento preferencial en la dirección (200).

Cuando el nivel de dopante se incrementa (región de bajo nivel) la intensidad para todos los picos de difracción excepto el (101) decrecen. Se infiere que una significativa contribución a la intensidad de los picos de difracción en rayos X proviene de la dispersión por los cristalitos dispersos en la superficie.

En el rango correspondiente a bajo nivel de dopante (0-8% F/Sn), los átomos de flúor contribuyen a disminuir la formación de cristalitos dispersos sobre la superficie de las películas, originando una superficie más lisa.

Cuando el nivel del dopante es alto (30% F/Sn), los picos (200) y (110) se incrementan, el pico (211) decrece ligeramente y el pico (101) queda prácticamente suprimido, (fig. 5d).

Estos resultados pueden ser comprendidos a partir de los estudios hechos por Fantini [2]. En efecto, se ha encontrado que cuando los átomos de flúor ocupan posiciones sustitucionales de oxígeno, el factor de estructura (200) y (101) decrece mientras que el factor de estructura (110) y (211) se incrementa. Para átomos de flúor en posiciones intersticiales el efecto es opuesto. Por lo tanto, a este nivel de dopaje, muchos de los átomos de flúor se encuentran en posiciones sustitucionales, por los cambios más significativos en los picos (110) y (101). Sin embargo, no se descartan posiciones intersticiales, aunque en menor proporción hecho que estaría indicado en el ligero incremento en el pico (200) y decrecimiento en el pico (211). A este nivel de concentración, se favorece la formación de una nueva fase; el Sn_3O_4 evidenciado por el pico en $2\theta = 31,5^\circ\text{C}$.

Finalmente se ha encontrado, para la región de alta concentración de dopante, una alta densidad de picaduras con un tamaño de $0,1\ \mu\text{m}$. Atribuimos este efecto a la formación de ácido fluorhídrico durante la deposición. En efecto, una elevada concentración de fluoruro de amonio favorece la formación del ácido produciéndose un ataque químico sobre la película. Esto puede constituir un resultado interesante, particularmente para ser usado en las aplicaciones como sensor de gas. La sensibilidad en la detección de gas mejora si la razón superficie/volumen se incrementa y el procedimiento usado en el presente trabajo ofrecería una manera sencilla de conseguirlo.

Conclusiones

La morfología y estructura de las películas de SnO_2 : F obtenidas en este trabajo, son altamente dependientes del contenido de flúor en la solución. Para bajos niveles de dopante, 0 - 8% F/Sn (peso) se obtienen superficies muy lisas. La transmitancia óptica y la conductividad eléctrica muestran un significativo mejoramiento con respecto a las muestras no dopadas. La apariencia de las películas iniciales mejora obteniendo luego alta transparencia.

Las películas dopadas con niveles por encima de 8% presentan picaduras en la superficie, esto causa un incremento en la resistividad. Se espera que estas películas mejoraran la sensibilidad en la detección de gas.

La difracción de rayos X muestra que para niveles de dopante de hasta 30% F, los átomos de flúor ocupan sitios sustitucionales de oxígeno. Sin embargo no se descarta la ocupación en posiciones intersticiales aunque en menor proporción.

REFERENCIAS

1. H. Haitjema and J. J. Ph. Elich, *Thin Solid Films* 205 (1993) 93.
2. M. Fantini and I. Torriani, *Thin Solid Films* 138 (1986) 255.
3. M. Miki and E. Andrade, *Thin Solid Films*, 224 (1993) 87.
4. J. Rickards, *Nucl. Instrum. and Meth.*, B56/57, (1991) 812.

Agradecimientos

Se aprecia el financiamiento proporcionado por: CONVENIO UNSA-TUD Holanda; IPPS Uppsala University Sweden; y la valiosa colaboración del DGLA de la Universidad Nacional Autónoma de México y al Instituto de Investigación de la Universidad Nacional de Ingeniería.

Divulgación especializada

Holografía

Carmen Eyzaguirre Gorvenia, Aníbal Valera Palacios (*)

RESUMEN

El presente trabajo describe los principios básicos de la Holografía, algunos aspectos de las técnicas actuales más importantes. Finalmente se describe el sistema Holográfico de geometría vertical desarrollado en la UNI y su aplicación a la elaboración de Hologramas de reflexión.

Introducción

La Holografía es un proceso fotográfico mediante el cual es posible producir imágenes tridimensionales. Las imágenes holográficas poseen todas las propiedades visuales del objeto reproducido, incluso el paralelaje entre sus partes (efecto de profundidad) con respecto al desplazamiento del observador. La observación es a simple vista y no requiere de lentes, ni de otros elementos.

El proceso holográfico fue inventado por Dennis Gabor (1948), quien le denominó holograma (del griego holos: totalidad) al registro fotográfico, indicando que se registra toda la información que genera la imagen, que de modo resumido, constituye un registro de ondas luminosas (amplitudes de la onda objeto), que son activadas de la placa, al iluminarse ésta con una luz de referencia.

Sin embargo, de la fecha en que Gabor elaboró el primer Holograma al presente, a pesar de las extraordinarias imágenes tridimensionales que se producen en profundi-

(*) Facultad de Ciencias de la UNI/ Feb. 1996, Laboratorio de Óptica.

dad (dentro de placas) o en alto relieve (flotando en el aire), esta técnica todavía dista mucho de haber logrado todas sus posibilidades.

¿QUÉ SE REQUIERE PARA HACER HOLOGRAMAS?

La obtención de un holograma (fotografía del objeto en tres dimensiones) ya sea de transmisión o reflexión requiere cumplir las siguientes consideraciones /ref 01/:

- Iluminación del objeto y la placa fotográfica (holograma) con un haz de luz coherente (igual longitud de onda, igual fase: Láser), gráfico 01.
- Debido a que en la placa holográfica interfieren los dos haces de luz (referencia/objeto), el sistema debe permanecer mecánicamente estable durante la exposición, para obtener un patrón de interferencia (holograma) definido.

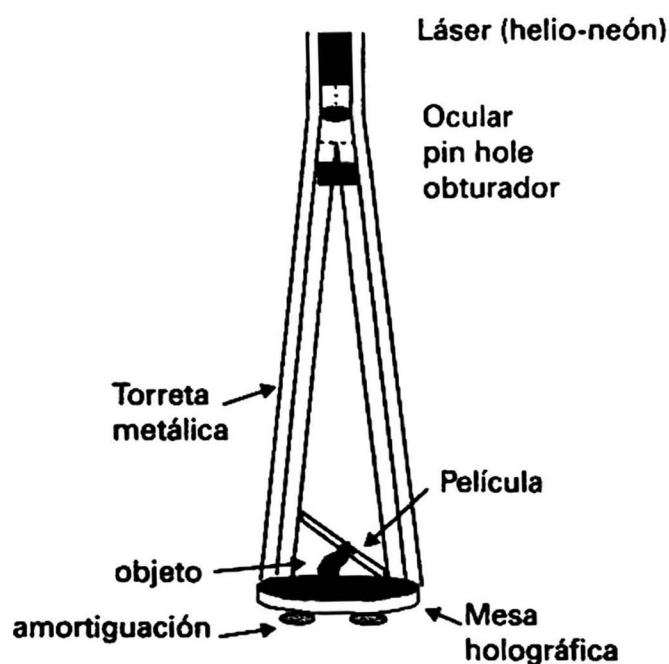


Figura 1: Diagrama del arreglo experimental Holográfico básico.

Para cumplir estas condiciones los métodos holográficos convencionales emplean:

- Láseres cw de potencia (15 mW a 50 mW).
- Mesas holográficas graníticas acopladas a muelles hidráulicos en configuraciones necesariamente horizontales.

El objetivo central de este trabajo fue construir y optimizar el modelo (UNI) de Cámara Holográfica de geometría vertical

Fundamento físico

La holografía se basa en los fenómenos de interferencia y de difracción.

INTERFERENCIA

a) Interferencia de dos ondas planas

Al hacer incidir simultáneamente dos ondas luminosas (planas) del mismo color (Láser) e intensidad sobre una pantalla desde dos direcciones ligeramente diferentes, la superficie de la pantalla parecerá a simple vista iluminada uniformemente. No obstante observando agudamente (o con ayuda de un microscopio) se podrá observar una serie de líneas, alternativamente brillantes y oscuras.

En algunos puntos de la pantalla las ondas están siempre en fase y se combinan dando lugar a las franjas brillantes, mientras que en otras las ondas están en oposición de fase, dando lugar a franjas oscuras.

El análisis del "patrón de interferencia" resultante, permite concluir:

- La separación entre las franjas de interferencia (D) depende del ángulo formado por las dos ondas luminosas y de la longitud de onda (Láser) correspondiente. Esta distancia es mayor cuanto menor es el ángulo de interferencia entre ellos.
- El contraste de las franjas (diferencia de luminosidad entre los máximos y mínimos de la luz) es máximo en el caso de que las intensidades de las ondas luminosas sean iguales, disminuyendo a medida que la diferencia de intensidades sea mayor.

El Holograma es una placa fotográfica en la que se registran las interferencias de las ondas que provienen de cada punto de un objeto con otra onda que se denomina de referencia (onda plana) por lo que se puede decir que un Holograma es el registro de un patrón de interferencias.

DIFRACCIÓN

Si el fenómeno de interferencia es la base de la obtención de un holograma, el responsable de la reconstrucción de la imagen holográfica es el fenómeno de difracción. En general se acostumbra tratar los fenómenos ópticos: Transmisión, reflexión y difusión como respuestas independientes. Sin embargo, en la teoría ondulatoria de la luz corresponden a diferentes aspectos de un mismo fenómeno llamado difracción.

a) Rejilla de difracción

Una rejilla de difracción se puede construir por el método holográfico, para lo cual procedemos a tomar una fotografía de la interferencia de dos ondas planas (efecto descrito anteriormente), al revelar la película la colocamos en la misma posición de registro y la iluminamos por una sola de las fuentes; en la observación aparecerán las dos fuentes: La real y la virtual.

Es de notar que la rejilla de difracción es un ejemplo simple de holograma.

HOLOGRAFÍA

Cuando se ilumina un objeto con luz monocromática, cada uno de los puntos del objeto se transforma en un emisor. La interferencia de una fuente de referencia (onda plana) con la multitud de pequeños emisores del objeto sobre la placa holográfica produce un patrón complejo (estructura de interferencia), la que al ser revelada y colocada en su posición original sujeta a la radiación de referencia, va a actuar como un elemento de difracción, reconstruyendo visualmente al objeto.

El problema de la holografía reside en la dificultad de obtener una onda de referencia que sea capaz de producir patrones de interferencia con la onda luminosa proveniente del objeto que sean a la vez estables y de buen contraste, al menos durante el tiempo que dura la exposición de la placa fotográfica. Si esto no se cumple, la calidad de la imagen será deficiente y de pobre luminosidad. En la época en que Gabor descubrió la holografía (1947), las fuentes luminosas disponibles tenían poca coherencia.

Gabor pudo obtener los primeros hologramas utilizando una lámpara de vapor de mercurio con un filtro que solo dejaba pasar la radiación correspondiente a un pequeño intervalo de longitudes de onda. Los objetos eran pequeñas transparencias (1 x 1 mm), de tal forma que la onda directamente transmitida actuaba como onda de referencia.

Uno de los inconvenientes es que con los medios disponibles los objetos debían ser muy sencillos y con unas características determinadas. Otro, que la luz difractada por el holograma que forma la imagen se propagaba en la misma dirección que la transmitida directamente lo que dificultaba enormemente su visualización. Los resultados no fueron por tanto muy espectaculares, lo que explica su letargo hasta 1962, fecha en que aparecieron los primeros Lasers comerciales, recibiendo Gabor en 1971 el premio Nobel.

Hologramas de Leith Upatnieks

Los científicos americanos E.N. Leith y J. Upatnieks (1962) fueron quienes, empleando un láser, propusieron el primer esquema práctico para la obtención de un holograma.

El haz emitido por un láser se divide en dos partes mediante una lámina semitransparente. Uno de estos nuevos haces, después de ser expandido por una lente sirve para iluminar el objeto, utilizando un espejo y la luz dispersada por este: haz del objeto, se recoge en una placa fotográfica. El otro haz, después de ser expandido es reflejado por otro espejo hacia la placa fotográfica: haz de referencia.

La reconstrucción se hace con un láser que emita luz del mismo color y de forma que la divergencia del haz sea la misma que la del haz de referencia en la obtención del holograma.

Hologramas de reflexión

En 1963, Y.N. Denisyuk (Rusia) y Van Heerden (USA) expandieron más aún los conceptos básicos desarrollados por Gabor, empleando el principio de Lippmann 1891, por medio del cual se emplea una emulsión fotosensible relativamente gruesa, permitiendo grabar planos de interferencias (Red), constituyendo el holograma en sí una red de difracción tridimensional con propiedades muy selectivas en la longitud de onda (filtro espectral de banda estrecha), efecto que permitió reconstruir los hologramas a partir de una fuente puntual de luz blanca.

CLASIFICACIÓN DE HOLOGRAMAS

Los hologramas se clasifican de acuerdo al tipo /Ref. 06/:

- Geometría de grabación.
- Modulación impuesta.
- Espesor del material fotosensible.
- Modo de la formación de imagen.

TEORIA BÁSICA

La holografía de sistema no axial es la mejor forma actual de realizar hologramas y es desde el punto de vista conceptual, el más simple de tratar teóricamente. En esta sección nos limitamos a describir los aspectos básicos de la teoría, según el caso de un holograma delgado.

Hologramas planos

Por un holograma plano (o delgado) se quiere dar a entender que se trata de un holograma cuya capa emulsiva fotosensible es delgada comparada con el espaciamiento de la frecuencia espacial más alta (separación mínima) del patrón de exposición, el cual es básicamente bidimensional.

Considerando la incidencia de dos frentes de onda a la placa holográfica, donde O es el frente de onda proveniente del objeto y R es el campo de referencia, ambos interfieren en la placa holográfica H.

El campo resultante en H es: $O + R$.

La irradiancia $I(x)$ sobre la placa fotográfica está dada por:

$$I(x) = (O + R)(O + R)^* = OO^* + RR^* + OR^* + O^*R$$

Asumiendo que estas intensidades de radiación corresponden a la región lineal de sensibilidad de la película, la amplitud de transmisión $T(x)$ en función de la resultante de exposición $E(x)$ será:

$$T(x) = T_o + B E(x)$$

Donde T_o es una constante (transmitancia de fondo) y B es un parámetro determinado por la sensibilidad de la emulsión y las condiciones del procesamiento. Puesto que $E(x)$ es directamente proporcional a $I(x)$, podemos escribir:

$$T(x) = T_o + B t (OO^* + RR^* + OR^* + O^*R)$$

donde t : es el tiempo de exposición.

Cuando un holograma con esta transmitancia es iluminada nuevamente con una onda de referencia R , la onda transmitida por el holograma es: $X(x)$

$$X(x) = R(x) T(x) = R T_o + B t (R O O^* + R R R^* + R R^* O + R O^* R)$$

Si la onda iluminante R es suficientemente uniforme sobre la placa, tal que RR^* toma un valor constante en todo el Holograma, el término RR^*O de la relación anterior resulta simplemente un valor constante $\times O$. Este término representa por tanto una onda que es idéntica a la onda del objeto original O , permitiendo su reconstrucción. El Holograma se comporta de este modo como una ventana, a través de la cual se observa una imagen virtual del objeto.

Los otros términos representan ondas "fantasmas":

- El término $ROO^* + RRR^*$ es la onda de orden cero.
- El término RRO^* es la onda de primer orden y es conocida como la onda conjugada.

A su vez, la imagen conjugada puede actuar como una imagen real o como una imagen virtual, dependiendo de la geometría particular usada para formar el holograma. Cuando la imagen es real, esta se encuentra espacialmente localizada en frente del holograma (del lado de la vista).

Estos dos últimos términos representan ondas visuales que perturban eventualmente la visualización del objeto, como era el caso en la experiencia de Gabor. En los arreglos actuales con el haz de referencia excéntrico, estas ondas están separadas espacialmente de la onda O objeto reconstruida. Esta es la mayor ventaja de la técnica "fuera de eje" y es la razón principal de su gran utilización.

Parte experimental

En lo que sigue describimos los elementos básicos, así como el método experimental seguido para la elaboración de hologramas.

ELEMENTOS ÓPTICOS UTILIZADOS PARA LA CÁMARA HOLOGRÁFICA

a) Banco óptico

El banco óptico juega un papel sumamente importante en la eliminación de vibraciones indeseables, por tal motivo se usan mesas graníticas sumamente masivas, tal que la frecuencia de resonancia sea suficientemente pequeña.

En nuestro caso se empleó una mesa de acero inoxidable de 50 kg y 18 cm de radio.

Sobre la mesa de acero se montó una estructura metálica de 1.8 m, en la que se dispuso la mayoría de los elementos ópticos del sistema.

b) Láser

El láser usado en la experiencia fue de Helio - Neón marca "Metrologic" de 8 mW de potencia. Posibilitando irradiancias netas sobre las placas, con máximo del orden de 0.1 W/m^2 .

En nuestro equipo el láser se dispuso en la parte superior de la torreta óptica, apuntando al centro de la mesa de acero.

c) Filtro espacial

El filtro espacial estaba constituido de un objetivo de microscopio y de un orificio "Pin hole" y tiene la doble función de expandir el haz láser y de otro lado eliminar los rayos secundarios creados por dispersión, las mismas que de no ser filtradas provocan inhomogeneidad en el plano de referencia. El diámetro del orificio fue de 50 μm (autoconstruido).

El sistema de filtro espacial se ajustó en la torreta inmediatamente por debajo del láser, tal que la mesa era iluminada homogéneamente.

d) Obturador

Debido a que las dosis de irradiancia láser recomendadas para las placas son del orden de: (100 J/cm^2), se hace necesario dosificar los tiempos de exposición, desde algunos segundos a minutos. Motivo por el cual se emplea obturadores electrónicos que ejecutan esta labor automáticamente. En nuestro caso usamos un obturador simple (cámara fotográfica) y el tiempo de exposición lo estimamos con un cronómetro.

La posición del obturador en la torreta fue a una altura intermedia por debajo del filtro espacial.

e) Espejos desviadores - Semi-espejos

En la tarea de grabación del holograma, se hace necesario hacer un ajuste mutuo de la intensidad de luz que incide tanto sobre el objeto, como así con el que llega directo a la placa (Referencia). Esta función se realiza convencionalmente con un espejo selectivo rotatorio: "Beam Splitter". En nuestro caso se tuvo que elaborar un juego de semiespejos (método de evaporación) para tal fin.

f) Portaplacas

El sistema portador de placas debe permitir de un lado ajustar debidamente las placas holográficas y de otro lado permitir un fácil manipuleo en completa oscuridad.

En nuestro caso se fabricó un portaplacas de armazón acrílico, con una corredera a presión donde es insertada la placa holográfica.

El Portaplacas se coloca sobre la mesa holográfica.

g) Placa holográfica

Debido a que en el patrón de interferencias que forman el holograma, las distancias medias son del orden de magnitud de la longitud de onda que se

emplea, se hace necesario usar una película de alta resolución para el acto de grabación: Placa holográfica. Las placas empleadas en holografía, comercialmente disponibles tienen resoluciones de 2000 a 2500 líneas por mm. Las placas utilizadas en este caso fueron de Eastman Kodak (649 F).

Un detalle importante en el montaje en el portaplaca es la ubicación correcta, según el lado sensible de la placa (localización de la emulsión), que debe estar dirigido en la dirección de incidencias de los rayos.

Otro detalle importante constituye el hecho que luego de instalar la placa holográfica en el fijador se debe esperar de 15 a 20 minutos para poder hacer la respectiva exposición, ya que las vibraciones producidas al instalar o ajustar tienen que desaparecer paulatinamente.

PROCEDIMIENTO PARA EL MONTAJE DEL SISTEMA

Los objetos a ser holografiados se colocan directamente sobre la mesa holográfica, o pegados sobre soportes acrílicos, para facilitar las tomas, adicionalmente se monta un sistema de espejos, tal que, tanto la placa como el objeto sean iluminados simultáneamente.

En el gráfico (02) se muestra el esquema final del arreglo global de la cámara holográfica.

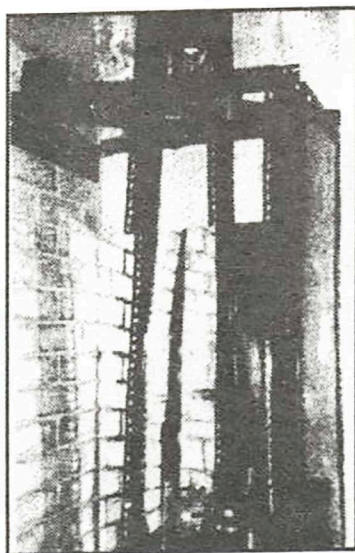


Figura 2. *Aparataje experimental Holográfico de geometría vertical, desarrollado en la UNI/ 1995.*

PROCEDIMIENTO EXPERIMENTAL DE OPTIMIZACIÓN SEGUIDO:

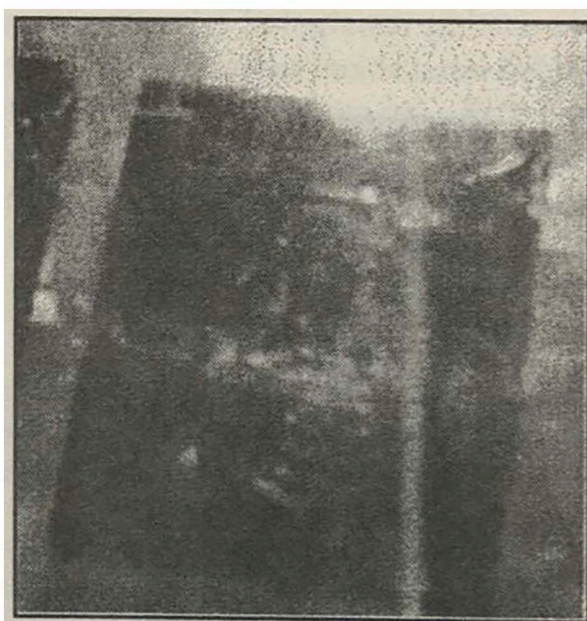
- a) Mantener constantes los parámetros relativos al procesamiento de la placa holográfica (revelado, fijado, lavado) según ref. /6/.
- b) Variar sistemáticamente los tiempos de exposición y niveles de iluminación relativa (placa holográfica, objeto).

Resultados

Los resultados de la optimización sistemática del sistema holográfico aplicado a la elaboración de hologramas de reflexión se presentan en un catálogo adjunto al trabajo (cerca de 250 muestras).

Las evaluaciones se realizaron teniendo en cuenta los siguientes criterios:

- Iluminación constante.
- Tiempo de exposición constante.
- Iluminación selectiva constante.



Conclusiones

Los resultados obtenidos son buenos, desde un punto de vista global, ya que las muestras obtenidas finalmente, luego de refinar la técnica permiten ya una buena visualización tridimensional del objeto, a partir de una fuente luminosa puntual (foco) o directamente al sol.

Es de observar el efecto crítico de la combinación de luces y tiempos de exposición total de la muestra.

- a) Para tiempos de exposición relativamente largos, la imagen toma un color verde oscuro (azulado), obteniéndose una imagen indefinida (muy oscura).
- b) Para tiempos de exposición relativamente cortos, la imagen toma un color rojizo pálido, obteniéndose una imagen indefinida (débil).
- c) Para tiempos de exposición correctos, la imagen toma un color verde brillante, resaltando nítidamente la imagen bajo un gran ángulo de observación.

No obstante los buenos resultados obtenidos, estos son aún factibles de mejorar, corrigiendo algunas limitaciones técnicas. Es de anotar que la mejor selección en la distribución de luces depende fuertemente de la muestra a reproducir.

En resumen:

Este trabajo ha demostrado la factibilidad de realizar reproducciones holográficas en base a un equipo simple desarrollado en la UNI para tal fin.

BIBLIOGRAFÍA

1. R. D. Bahuguna, D. Malacara *"Holography", Methods of experimental Physics. Vol. 26, Academic Press, USA 1988.*
2. H. J. Tung *"The one beam transmission hologram". The physics teacher, USA. Feb. 1980 (129 - 133).*
3. H. M. Smith *"Holographic recording materials". Topics y applied Physics, 1985, Vol 20 USA.*
4. M. Francon *"Holografia". Ed. Paraninfo, Madrid, 1972.*
5. F. Unterseher, B. Schlesinger, J. Harrison *"Holography Handbook: making holograms the easy way". Ross Books, N. Y. (1982).*
6. C. Eyzaguirre *"Desarrollo de un sistema holográfico de geometría vertical". Tesis de Licenciatura, Febrero 1996, UNI.*
7. H. Kogelnick *Bell Syst. Tech. J. 48, 2909 (1969).*

El vector de inercia

Holguer G. Valqui (*)

El estudio de los momentos de inercia y los ejes principales de inercia se realiza usualmente por medio de la llamada matriz de inercia. En el presente artículo se construye el vector de inercia, el mismo que permite un análisis "más visible" o más geométrico de los mencionados conceptos, en los cuales sólo aparecen los vectores y sus productos.

Sea un eje (Q, u) ; que pasa por el punto Q , en la dirección y sentido del vector u , unitario. Sea una partícula puntual de masa m situada en el punto P . Definamos el vector de inercia $Y(Q, u)$ de dicha partícula con respecto al eje (Q, u) :

$$Y(Q, u) = m \cdot QP \times (u \times QP) \quad [01]$$

Por conveniencia supondremos que el vector de inercia está aplicado en el punto Q . De la definición podemos apreciar que $Y(Q, u)$ es un vector que se encuentra en el plano definido por el eje y el punto P , siendo perpendicular a QP .

Definamos también el momento de inercia $I(Q, u)$, de dicha partícula con relación al citado eje, como la proyección de Y sobre u :

$$I(Q, u) = \langle u, Y(Q, u) \rangle = m \cdot (u \times QP)^2 \geq 0 \quad [02]$$

(*) Facultad de Ciencias, Universidad Nacional de Ingeniería, Casilla postal 31-139, Lima - Perú.

[donde hemos usado la propiedad del producto triple, según la cual, $\langle u, v \times w \rangle = \langle u \times v, w \rangle$, y los corchetes angulares expresan el producto escalar]. Es decir, dicho momento es el producto de la masa (puntual) por el cuadrado de su distancia al eje.

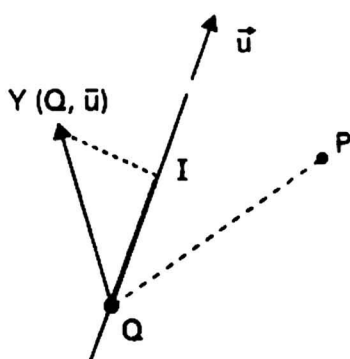


Fig. 1 $Y(Q, u)$ para una masa en P

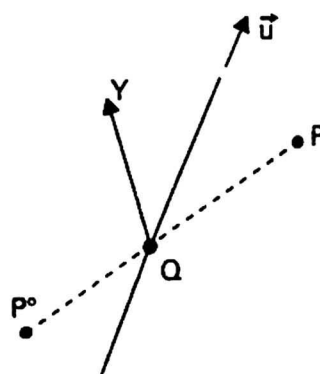


Fig. 2 $Y(Q, u)$ para m en P^c

Consideremos ahora el eje (Q, u) y una distribución (que aquí supondremos discreta) de partículas puntuales de masa m_k situadas en las posiciones P_k . Entonces, con respecto a tal eje, podemos definir el vector de inercia del conjunto, como la suma de los vectores de inercia de cada una de las partículas,

$$Y(Q, u) = \sum_k m_k \cdot QP_k \times (u \times QP_k) \quad [03]$$

y el correspondiente momento de inercia, como la proyección de dicho vector sobre la recta del eje:

$$I(Q, u) = \sum_k m_k \cdot (u \times QP_k)^2 \geq 0 \quad [04]$$

lo cual coincide con la definición usual de momento de inercia. Nótese que $|u \times QP_k|$ es la distancia del punto P_k a la recta (Q, u) y que, en general, el vector de inercia no cae sobre el eje.

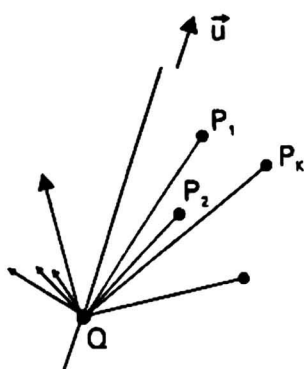


Fig. 3 $Y(Q, u)$

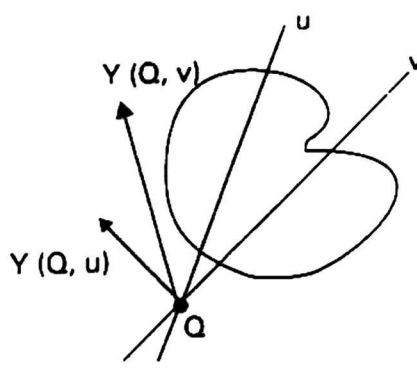


Fig. 4 $Y(Q, v)$

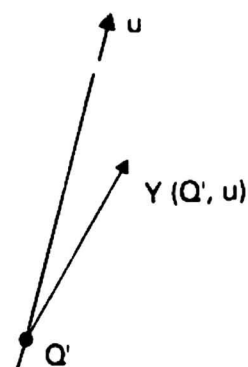


Fig. 5 $Y(Q', u)$

Ejes principales de inercia

Dado un punto Q , diremos que un eje (Q, u) es eje principal de inercia de una distribución de partículas, si el correspondiente vector de inercia es paralelo al eje (o, si se prefiere, cae sobre el eje), es decir,

$$(Q, u) \text{ es EPdI} \Leftrightarrow Y(Q, u) = \beta u \quad [05]$$

donde, como se verifica fácilmente, $\beta = I(Q, u)$.

Si ahora consideramos dos ejes cualesquiera (Q, u) y (Q, v) , podemos verificar que,

$$\langle v, Y(u) \rangle = \langle u, Y(v) \rangle \quad [06]$$

Si dichos ejes fuesen ejes principales obtendremos que

$$[I(u) - I(v)] \langle u, v \rangle = 0,$$

de donde podemos deducir que, si los momentos de inercia correspondientes a dichos ejes son diferentes, entonces los ejes deben ser perpendiculares entre sí.

Por otra parte, si (Q, u) y (Q, v) son dos ejes principales de inercia, a los cuales corresponde un mismo momento I , entonces cualquier eje (Q, w) , situado en el plano determinado por los ejes anteriores, también será un eje principal, con el mismo valor propio I . Es decir, si $Y(Q, u) = Iu$, $Y(Q, v) = Iv$, entonces $Y(Q, w) = Iw$, donde w es un vector unitario, combinación lineal de u y v .

Es claro que lo anterior también se cumple para el caso de tener tres ejes principales con el mismo I .

Sean ahora dos partículas de igual masa m , situadas simétricamente con respecto a un eje (Q, u) , en los puntos P y P' . Es decir, PP' es perpendicular al eje, o $\langle u, PP' \rangle = 0$.

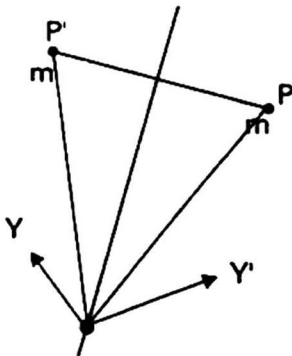


Fig. 6 Y es para m en P . Y' es para m en P'

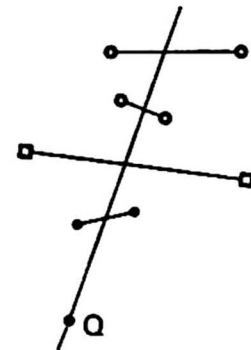


Fig. 7 Un cuerpo formado por pares de "partículas", simétricas dos a dos.

Entonces, si B es la intersección de PP' con el eje, podemos escribir,

$$QP = QB + BP, \quad QP' = QB - BP$$

$$u \times QP = u \times BP, \quad u \times QP' = -u \times BP$$

$$QP \times (u \times QP) + QP' \times (u \times QP') = QP \times (u \times BP) - QP' \times (u \times BP) =$$

$$(QP - QP') \times (u \times BP) = P'P \times (u \times BP) = 2BP \times (u \times BP) = 2(BP)^2 u$$

es decir, el vector de inercia para dichas dos partículas será

$$Y(Q, u) = 2m(BP)^2 u, \quad [07]$$

de donde resulta que (Q, u) es un eje principal de inercia para el sistema constituido por dichas dos partículas.

Sea $\{(m_k, P_k) / k = 1 \dots 2n\}$ un sistema de $2n$ partículas, de manera que ellas posean, dos a dos, la misma masa y sean simétricas con respecto al eje; es decir, por ejemplo, $m_k = m_{k'}$, $\langle u, P_k P_{k'} \rangle = 0$, con $k' = n+k$, para $k = 1 \dots n$. Entonces, (Q, u) será eje principal para cada par de partículas; entonces será también un eje principal para todo el sistema de las $2n$ partículas. En particular, todo cuerpo homogéneo, con simetría poligonal par, con respecto a cierta recta, tendrá a dicha recta como eje principal de inercia: dicho cuerpo puede ser considerado como la suma de pares de "pedacitos" de igual masa, simétricos con relación a la recta considerada. Un caso más particular lo constituye un cuerpo con simetría rotacional alrededor de la citada recta.

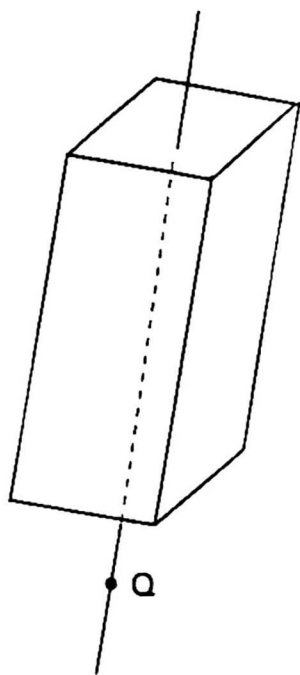


Fig. 8 Un cuerpo con simetría par con respecto a (Q, u).

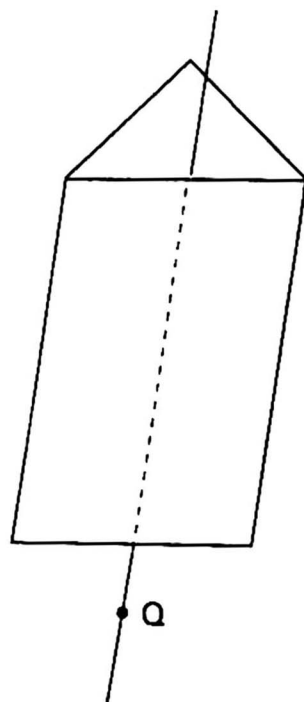


Fig. 9 Un cuerpo con simetría impar con respecto a (Q, u)

Sea B un punto del eje (Q, u) y sean r masas puntuales de igual magnitud que se encuentran en los puntos P_1, P_2, \dots, P_r de un plano que pasando por B es perpendicular al eje, de manera que la suma de los vectores $BP_1 + BP_2 + \dots + BP_r$ sea nula. Entonces tendremos que

$$QP_k \times (u \times QP_k) = (QB + BP_k) \times [u \times (QB + BP_k)] = -\langle QB, u \rangle BP_k + (BP_k)^2 u$$

de donde obtenemos que el eje (Q, u) resulta ser eje principal de inercia para tal configuración de r partículas de masas iguales. A partir de esto se puede demostrar que, si un cuerpo homogéneo tiene secciones transversales (perpendiculares al eje) poligonales, que cumplan la condición mencionada anteriormente, entonces dicho eje resulta ser eje principal de inercia.

Notemos, por otra parte, que $(-QP) \times [u \times (-QP)] = QP \times (u \times QP)$. Es decir, si P° es simétrico de P con respecto a Q, $QP^\circ = -QP$, entonces el vector de inercia de una partícula en P° será el mismo que si la partícula estuviese en P.

Sea P'' un punto simétrico de P con respecto al plano que pasando por Q es perpendicular al eje (Q, u) . Entonces, el punto P' , tal que $QP' = -QP''$, será simétrico al punto P, con respecto al eje. Pero tal eje es eje principal para las dos masas puntuales (y de igual magnitud) situadas en P y en P' ; consecuentemente, (Q, u) será eje principal para las dos masas puntuales situadas en los puntos P y P'' .

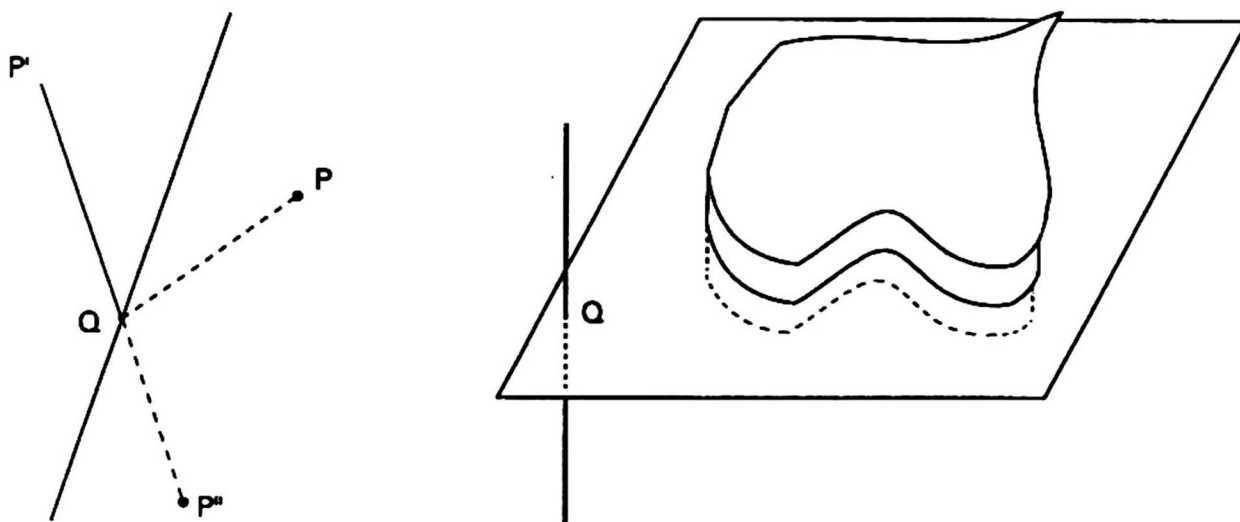


Fig. 10 Los puntos P y P' son simétricos con respecto a (Q, u) ; P' y P'' son simétricos con respecto a Q, resultando que P y P'' son simétricos con respecto al plano perpendicular al eje.

Sea ahora $\{(m_k, P_k) / k=1 \dots 2n\}$ un sistema de $2n$ partículas de manera que ellas tengan, dos a dos, la misma masa y sean simétricas con respecto a un plano que, pasando por Q, es perpendicular al eje, entonces dicho eje, (Q, u) , resultará ser eje

principal de inercia de todo el sistema de las $2n$ partículas. En particular, si un cuerpo homogéneo tiene un plano de simetría, todo eje perpendicular a dicho plano resulta ser eje principal de inercia para dicho cuerpo.

Sean e_1, e_2, e_3 , vectores unitarios no paralelos a un mismo plano (por ejemplo, en las direcciones de los ejes de coordenadas). Sea $Y_k = Y(Q, e_k)$, para $k = 1, 2, 3$. Entonces, para el vector de inercia $Y(Q, u)$, donde $u = \beta_1 e_1 + \beta_2 e_2 + \beta_3 e_3$ tendremos que,

$$Y(Q, u) = \beta_1 Y_1 + \beta_2 Y_2 + \beta_3 Y_3 \quad [08]$$

lo cual, conocidos los vectores de inercia con respecto a los ejes coordenados (traslados al punto Q), nos permite calcular el vector de inercia con respecto a cualquier eje que pase por Q .

[Notemos que por ser los tres vectores no paralelos a un mismo plano, entonces cualquier otro vector puede ser expresado como combinación lineal de ellos].

Determinación de los ejes principales

[En lo que sigue, como todo se va a referir a un mismo punto Q , dejaremos de escribir tal punto]. Como ya hemos visto, la ecuación que satisface un eje principal es $Y(u) = I(u) \cdot u$, de manera que para el producto con cualquiera de los vectores unitarios según los ejes de coordenadas tendremos

$$\begin{aligned} \langle e_k, Y(u) \rangle &= I(u) \langle e_k, u \rangle && \text{o, teniendo presente [06],} \\ \langle u, Y_k \rangle &= I(u) \langle u, e_k \rangle, && \text{es decir,} \\ \langle u, Y_k - I(u) e_k \rangle &= 0, && \text{para } k = 1, 2, 3 \end{aligned} \quad [09]$$

Como el vector u es no nulo, entonces los tres vectores $Y_k - I(u) e_k$ deberán ser coplanares, es decir, su producto triple debe ser nulo,

$$\langle Y_1 - I(u) e_1, [Y_2 - I(u) e_2] \times [Y_3 - I(u) e_3] \rangle = 0 \quad [10]$$

que es una ecuación cúbica para $I(u)$:

$$I(u)^3 - A I(u)^2 + B I(u) - C = 0 \quad [11]$$

con

$$A = I(e_1) + I(e_2) + I(e_3), \quad C = \langle Y_1, Y_2 \times Y_3 \rangle$$

$$B = \langle e_1, Y_2 \times Y_3 \rangle + \langle e_2, Y_3 \times Y_1 \rangle + \langle e_3, Y_1 \times Y_2 \rangle$$

Resolviendo la ecuación [11] se obtendrán tres valores posibles para $I(u) = r_1, r_2, r_3$. Esos tres valores, o raíces, serán números positivos (reales), como se ve en [04]. Cada raíz corresponderá a un cierto vector unitario, de manera que, en general, tendremos tres vectores u_1, u_2, u_3 tales que $I(u_k) = r_k, k = 1, 2, 3$.

Notemos que la igualdad [09] se satisface para cada uno de los vectores u_1, u_2, u_3 . Es decir, por ejemplo, u_1 es perpendicular a cada uno de los vectores $Y_1 - r_1 e_1, Y_2 - r_1 e_2, Y_3 - r_1 e_3$, los mismos que son coplanares (en rigor, paralelos a un mismo plano). Entonces pueden presentarse tres situaciones:

- Los tres vectores son independientes dos a dos. En tal caso, el vector u_1 será paralelo al producto vectorial de dos de ellos (que lo dividiremos entre su módulo para obtener un vector unitario),
- Dos de los vectores son linealmente independientes; el tercero es paralelo a uno de los anteriores. En este caso, u_1 será paralelo al producto vectorial de los dos primeros,
- Los tres vectores son paralelos entre sí; en este caso u_1 será cualquier vector unitario perpendicular a la dirección de los vectores anteriores.

Simetría y degeneración

Aquí queremos destacar la relación que existe entre la simetría geométrica de un cuerpo y la degeneración de las raíces de la ecuación [11].

- El cilindro homogéneo.-** Como hemos mostrado anteriormente, el eje (Q, u_3) de un cilindro homogéneo (por poseer simetría de revolución alrededor de dicho eje) es un eje principal para dicho cilindro, con momento de inercia $I(u_3)$. Si los otros momentos de inercia, $I(u_1)$ y $I(u_2)$, son diferentes a $I(u_3)$, entonces por [06], tanto u_1 como u_2 serán perpendiculares a u_3 . Es decir, u_1 y u_2 deberán estar sobre el plano perpendicular a u_3 . Pero, debido a la simetría cilíndrica del cuerpo en consideración, no existe manera de distinguir la dirección u_1 de la u_2 ; es decir, tanto u_1 como u_2 son direcciones cualesquiera, perpendiculares a u_3 y, consecuentemente, toda dirección en el plano perpendicular a u_3 , será la de un eje principal. Nótese que la situación recíproca no tiene porque ser cierta: un cuerpo (o un sistema de partículas) puede tener dos momentos de inercia iguales, sin que dicho cuerpo posea simetría cilíndrica.

Cuando un cuerpo posea dos momentos de inercia iguales, diremos que él posee simetría cilíndrica dinámica con respecto al eje en consideración. Entonces un cuerpo homogéneo con simetría cilíndrica geométrica poseerá también simetría cilíndrica dinámica; pero no recíprocamente.

- II. *La esfera homogénea.*- Sea una esfera homogénea con centro en Q; entonces ella será simétrica alrededor de un eje (Q, u), con lo cual sus momentos de inercia con respecto a los ejes perpendiculares a u serán todos iguales entre sí. Pero u es una dirección arbitraria (pues en una esfera homogénea no hay manera de distinguir una dirección de otra); es decir, los momentos de inercia en cualquier dirección son iguales, o las raíces de la ecuación [11] son coincidentes (una raíz triplemente degenerada). Aquí también debemos observar que si para cierto cuerpo se obtienen las tres raíces de [11] iguales entre sí, ello no significa que se trate de un cuerpo con simetría esférica. Por ejemplo, podría tratarse de un cilindro, como hemos visto anteriormente, en el que las tres (y no sólo dos) raíces coinciden.

Cuando para un cierto cuerpo resulten iguales sus tres momentos principales de inercia, diremos que dicho cuerpo tiene, con respecto al eje considerado, simetría dinámica esférica.

Simetrización dinámica

Consideremos el siguiente ejemplo: Sea Q el origen de un sistema cartesiano. Coloquemos, tanto en punto (a, 0, 0) como en (-a, 0, 0) una masa puntual m_1 ; coloquemos, tanto en (0, b, 0) como en (0, -b, 0) una masa puntual m_2 , y tanto en (0, 0, c) como en (0, 0, -c) una masa puntual m_3 . Entonces, podemos verificar directamente que para tal sistema de 6 partículas se cumple que,

$$Y(e_1) = (m_2 \cdot b^2 + m_3 \cdot c^2) e_1$$

$$Y(e_2) = (m_3 \cdot c^2 + m_1 \cdot a^2) e_2$$

$$Y(e_3) = (m_1 \cdot a^2 + m_2 \cdot b^2) e_3$$

con lo cual los tres ejes coordenados resultan ser ejes principales de inercia. Si ahora exigimos $m_2 \cdot b^2 = m_1 \cdot a^2$, tendremos que el sistema de seis partículas poseerá simetría dinámica cilíndrica. Si, además, exigimos que $m_3 \cdot c^2 = m_1 \cdot a^2$, entonces dicho sistema poseerá simetría dinámica esférica.

Consideremos ahora un cuerpo cualquiera que, según tres ejes perpendiculares entre sí, posea los momentos de inercia I_1, I_2, I_3 . Coloquemos, como en el ejemplo anterior, las seis masas puntuales m_1, m_2, m_3 . Entonces los nuevos momentos de inercia serán $I_1' = I_1 + (m_2 \cdot b^2 + m_3 \cdot c^2)$, etc. Es decir, eligiendo las masas y las distancias adecuadamente, podemos lograr que el cuerpo dado adquiriera simetría dinámica cilíndrica o esférica.

Notemos que, en lo anterior, lo dicho sobre el vector de inercia ha sido independiente de que las partículas consideradas se encuentren en movimiento o no. A continuación veremos el papel que desempeña el vector de inercia en el movimiento de rotación de un cuerpo rígido.

El momento angular de un cuerpo rígido

Sea un cuerpo rígido que, en el instante t rota alrededor de un eje (Q, u) con velocidad angular $w \cdot u$, siendo Q un punto fijo al cuerpo. La velocidad p° , con la que se mueve un punto P del cuerpo, con respecto a una cierta Referencia será

$$p^\circ = q^\circ + w \cdot uxQP \quad [12]$$

donde q° es la velocidad de Q con respecto a dicha referencia.

El momento angular L_Q del cuerpo rígido (que aquí lo consideraremos como un conjunto de partículas) con respecto al punto Q , está dado por

$$\begin{aligned} L_Q &= \sum QP_k \times (m_k \cdot p_k^\circ) = \sum m_k QP_k \times (q^\circ + dQP_k/dt) = \\ &= \sum m_k QP_k \times (q^\circ + w \cdot uxQP_k) = (\sum m_k QP_k) \times q^\circ + w \cdot \sum m_k QP_k (uxQP_k) \end{aligned} \quad [13]$$

Si (Q, u_1) , (Q, u_2) y (Q, u_3) son tres ejes principales que pasan por Q , perpendiculares entre sí, con los correspondientes momentos de inercia I_1, I_2, I_3 , entonces tendremos

$$u = \beta_1 u_1 + \beta_2 u_2 + \beta_3 u_3$$

y, teniendo presente [08]

$$Y(Q, u) = I_1 \beta_1 u_1 + I_2 \beta_2 u_2 + I_3 \beta_3 u_3$$

A su vez, la velocidad angular, $w \cdot u$, puede ser descompuesta en las tres direcciones principales

$$w \cdot u = w_1 u_1 + w_2 u_2 + w_3 u_3 \quad [14]$$

$$\text{de donde resulta que } w_k = \beta_k w \quad [15]$$

lo cual nos permite escribir, reemplazando en [13]

$$L_Q = m QG \times q^\circ + (I_1 w_1 u_1 + I_2 w_2 u_2 + I_3 w_3 u_3) \quad [16]$$

donde debe tenerse presente que los ejes principales del cuerpo rígido permanecieran ligados al cuerpo, es decir, los tres momentos de inercia no se modificarán al pasar el tiempo. En cambio, las direcciones u_1, u_2, u_3 , si variarán, cumpliéndose

$$u_k^\circ = (w.u) x u_k, \quad k = 1, 2, 3 \quad [17]$$

lo cual, en cada uno de los tres casos da

$$u_1^\circ = w_3 u_2 - w_2 u_3, \quad u_2^\circ = w_1 u_3 - w_3 u_1, \quad u_3^\circ = w_2 u_1 - w_1 u_2 \quad [17a]$$

Si ahora definimos

$$U_1 = I_1 \cdot w_1^\circ - (I_2 - I_3) w_2 w_3, \quad U_2 = I_2 \cdot w_2^\circ - (I_3 - I_1) w_3 w_1, \quad U_3 = I_3 w_3^\circ - (I_1 - I_2) w_1 w_2 \quad [18]$$

tendremos que

$$L_Q^\circ = m QGxq^{\circ\circ} + m (QG)^\circ x q^\circ + U_1 u_1 + U_2 u_2 + U_3 u_3 \quad [19]$$

Por otra parte, si consideramos el torque M_Q , con respecto al punto Q, de las fuerzas que actúan sobre el cuerpo rígido en consideración, tendremos que en un sistema inercial (terrestre)

$$M_Q = \sum QP_k x (m_k \cdot p_k^\circ)^\circ + \sum (QP_k)^\circ x (m_k \cdot p_k^\circ) - \sum (p_k^\circ - q^\circ) x (m_k \cdot p_k^\circ)$$

donde los dos últimos sumandos se anulan. De la suma de los dos primeros obtenemos la derivada del momento angular, es decir,

$$M_Q = L_Q^\circ - m \cdot q^\circ x g^\circ \quad [20]$$

donde g° es la velocidad del centro de masa. Es decir, teniendo presente que $(QG)^\circ x q^\circ = (g^\circ - q^\circ) x q^\circ = g^\circ x q^\circ$,

$$M_Q = m QGxq^{\circ\circ} - \sum U_k u_k \quad [21]$$

que es la ecuación de Euler para un cuerpo rígido

BIBLIOGRAFÍA

- H. G. Valqui, *Apuntes de Física General 1*, UN 1972.
H. Goldstein, *Classical Mechanics*, Addison-Wesley 1950.

Optimización usando el método del recocido simulado

René Víctor Valqui Vidal ()*

RESUMEN

En el presente trabajo se expresan las ideas fundamentales del método de optimización conocido como recocido simulado (inglés: simulated annealing) y se presentan además guías para la construcción de este tipo de algoritmos. Finalmente, diferentes aplicaciones serán brevemente expuestas.

Introducción

Recocido simulado (RS) es un método estocástico de búsqueda muy poderoso que se ha aplicado para resolver problemas en una variedad de disciplinas. Por ejemplo en física, matemática, ingeniería, optimización, investigación operacional, estadística, etc. En esta nota estaremos estudiando la aplicación del método RS a los problemas de optimización combinatoria.

Los problemas de optimización combinatoria se describen como:

(*) Instituto de Modelamiento Matemático, Universidad Técnica de Dinamarca, DK-2800 Lyngby.

Dado un espacio de configuración finito $S = \{x \mid x = (x_1, x_2, \dots, x_m)\}$, donde m es la dimensión del espacio, y una función costo $C: S \rightarrow \mathbb{R}$, que asigna un número real a cada configuración, se desea encontrar una configuración minimal $x^* \in S$, de manera que

$$\forall y \in S, C(x^*) \leq C(y)$$

Durante los últimos años ha habido un interés muy grande en RS porque se ha demostrado que muchos problemas de optimización combinatoria se han resuelto en un tiempo de computación muy razonable. Muchos de estos problemas son conocidos como NP-completos: todas las técnicas conocidas para obtener una solución exacta demandan un crecimiento exponencial en el número de iteraciones como función del tamaño del problema. Por esta razón se han desarrollado métodos heurísticos, como RS, para resolver este tipo de problemas. La diferencia entre un método heurístico y un método exacto es que el primero no garantiza una solución óptima; él es diseñado para dar una solución aceptable en un tiempo aceptable. RS no es un algoritmo en el sentido que da una secuencia mecánica de computación para resolver el problema, como es el caso del método simplex para resolver problemas de programación lineal. RS es una **estrategia** para resolver problemas de optimización; varias decisiones se deben tomar basadas en experimentos intensivos con la computadora para finalmente diseñar un algoritmo para resolver un problema particular. Ha habido un interés muy grande en la aplicación de RS a problemas de optimización combinatoria después de la publicación de los trabajos de Kirkpatrick et al (1983) y Černý (1985). Bibliografías se pueden encontrar en Collins et al (1988) y Koulamas et al (1994). Libros sobre RS han sido recientemente publicados, por ejemplo Azencott (1992) y Valqui Vidal (1993).

El objetivo de esta nota es de dar una presentación básica para entender y rápidamente aplicar RS para resolver problemas prácticos.

La idea de RS se basa en la analogía entre el problema de encontrar estados de energía mínima en un sistema físico y el problema de encontrar una configuración de costo mínimo en un problema de optimización combinatoria. Para resolver el problema primero, el material es recocido, o sea calentado a una temperatura que permite muchos rearrreglos atómicos, para después enfriarlo con cuidado, lentamente, hasta que el material esté "helado" en un estado de energía mínima. La técnica del RS utiliza un conjunto de "enfriamientos controlados" en un problema de optimización no físico, transformando una solución pobre y desordenada en otra solución optimizada.

Los algoritmos básicos

En esta sección mostraremos que el proceso de diseño de un algoritmo RS es fácil y se puede hacer muy rápidamente. Además, RS es un método **transparente**, eso

significa que personas que no son especialistas en optimización pueden comprender fácilmente sus principios. Probablemente, la manera más fácil de introducir RS es mostrando que es una modificación del método tradicional conocido como la heurística de la búsqueda local (inglés: local search heuristics).

La heurística de la búsqueda local

Una estrategia de búsqueda local comienza con una solución inicial $x_1 \in S$ y a cada iteración i , una nueva solución x_{i+1} es escogida en la vecindad $N(x_i)$ de la solución actual x_i .

La estrategia de búsqueda local más conocida es el algoritmo conocido como "descendimiento a pique" (inglés: steepest descent). Este algoritmo es mostrado en la tabla N° 1. Acá, una solución nueva x_{i+1} es encontrada tomando la mejor en $N(x_i)$ eso es una solución $x_{i+1} \in N(x_i)$ de manera que

$$C(x_{i+1}) \leq C(x_i), \quad \forall x \in N(x_i)$$

Luego, x_{i+1} será la nueva solución actual si no es peor que x_i .

PROCEDURE ALGORITMO DESCENDIMIENTO

Comenzar

Inicializar ($x_i, i = 1$)

Repetir

Encontrar mejor $\bar{x} \in N(x_i)$

Calcular $\Delta C = C(x_i) - C(\bar{x})$

Si $\Delta C \leq 0$, luego $x_i = \bar{x}$

Si no: pare

Fin

Tabla N° 1: Una heurística de búsqueda local

Esta estrategia parece muy razonable, es muy simple su implementación y es muy rápida su ejecución. Pero tiene un problema muy serio: es fácilmente atrapado

en **mínimo local**: soluciones buenas en una vecindad pequeña pero no necesariamente el mínimo global.

El algoritmo RS

RS ofrece una estrategia muy similar al algoritmo de descendimiento a pique con una diferencia principal: RS permite perturbaciones moviéndose en sentido ascendente de una manera controlada. Esto permite encontrar una forma de alejarse de un mínimo local.

RS trabaja de acuerdo al principio siguiente: al comienzo (temperatura alta) casi todos los movimientos (esto es las soluciones generadas de una solución actual, generadas aleatoriamente en su vecindad) son aceptados. Esto permite la exploración de S. Luego, gradualmente, se baja el parámetro de control (temperatura) con el objetivo que el algoritmo sea más y más selectivo aceptando nuevas soluciones generadas. Al final, solamente los movimientos que decrecen C son aceptados.

En RS la probabilidad de aceptar un movimiento que produce un incremento, ΔC , en la función de costo es llamada la **función de aceptación**:

$$\exp \left(-\frac{\Delta C}{T} \right)$$

donde T es el parámetro de control que corresponde a la temperatura en la analogía con el proceso físico de recocimiento. En el algoritmo RS, al inicio T tiene un valor muy alto para evitar ser atrapado prematuramente en un mínimo local. Luego, T es disminuido gradualmente hasta que se acerca a cero. Al final, la configuración "helada" es tomada como la mejor solución.

El algoritmo RS es presentado en dos versiones diferentes en las tablas N° 2 y N° 3.

PROCEDURE RS ALGORITMO INHOMOGÉNEO

Comenzar

Iniciación (k, T_k , i);

Repetir

Generar configuración (i \leftrightarrow j)

Calcular $\Delta C_{ij} = C(i) - C(j)$

Si $\Delta C_{ij} \leq 0$ luego $i := j$; si no
 Si $\exp(-\Delta_{ij} / T_k) > \text{aleatoria}[0,1]$ luego $i := j$;
 $k := k+1$
 Recalcular (T_k) ;
 Hasta parar;
 Fin;

Tabla Nº 2: RS algoritmo inhomogéneo

PROCEDURE RS *ALGORITMO HOMOGÉNEO*

Comenzar

Iniciación (k, T_k, L_k) ;

Repetir

Para $i := 1$ a L_k hacer

Comenzar

Generar configuración $(i \leftrightarrow j)$;

Calcular $\Delta C_{ij} = C(i) - C(j)$

Si $\Delta C_{ij} \leq 0$ luego $i := j$; si no

Si $\exp(-\Delta_{ij} / T_k) > \text{aleatoria}[0,1]$ luego $i := j$;

Fin

$k := k+1$;

Recalcular (L_k)

Recalcular (T_k)

Hasta parar;

Fin;

Tabla Nº 3: RS algoritmo homogéneo

Estas dos versiones diferentes de algoritmo RS representan dos modos diferentes de disminuir la temperatura T ; este proceso es denominado la estrategia de enfriamiento o el esquema de recocido (inglés: cooling strategy or annealing scheme):

- i. El algoritmo **inhomogéneo**, donde T disminuye después de cada movimiento, de esta manera el proceso se puede describir teóricamente como una cadena simple de Markov inhomogénea, y
- ii. El algoritmo **homogéneo**, donde T disminuye después de un número de movimientos L ; de esta manera el proceso se puede describir teóricamente como una secuencia de cadenas de Markov homogéneas, cada una generada a una temperatura fija T .

El análisis matemático de los procesos estocásticos generados por estos algoritmos dan las condiciones necesarias y suficientes para asegurar convergencia asintótica de los algoritmos hacia la solución óptima global con probabilidad uno (Aarts and Korst (1989)). Desgraciadamente estas condiciones no se pueden satisfacer en tiempo finito. Por eso, se tiene que especificar primero un número **finito** de movimientos para cada T , y después una secuencia **finita** de valores de T . Esto significa que en la práctica los algoritmos RS son métodos heurísticos que producen buenas soluciones que pueden ser las globales.

La implementación

Para resolver un problema particular de optimización combinatoria usando el método del RS, varias decisiones se tienen que tomar para diseñar un algoritmo apropiado. En la tabla N° 4 estas decisiones son presentadas. Primero, tenemos las decisiones **genéricas** que también se denominan como el esquema de enfriamiento. Segundo, tenemos las decisiones **específicas del problema**.

DECISIONES	
Genéricas (enfriamiento)	Específicas
<ul style="list-style-type: none"> • T_0 (temperatura inicial) • L_k (número de iteraciones) • T_k (función de temperatura) • Criterio de paro 	<ul style="list-style-type: none"> • i_0 (solución inicial) • generación de una solución vecina. • Cálculo de ΔC_{ij}

Tabla N° 4: Diseño de RS algoritmos

Decisiones genéricas

Una de las funciones de temperatura más usadas es la llamada esquema de enfriamiento geométrico.

$$T_{k+1} = \alpha T_k, \quad 0 < \alpha < 1$$

En la práctica buenos resultados se obtienen cuando, $0,8 \leq \alpha \leq 0,99$.

Cuando se está implementando el algoritmo inhomogéneo los parámetros que se deben fijar son T_0 , T_f (temperatura final) y α . En la versión homogénea se tiene que fijar también L_k . Generalmente en la práctica $L_k = L, \forall k$, y el valor de L como los valores de los otros parámetros se fijan por experimentación en la computadora.

Existen muchos otros esquemas de enfriamiento; para una presentación completa vease Collins et al (1988). La práctica muestra que lo importante es la velocidad de enfriamiento y no tanto la forma de este proceso.

La temperatura inicial T_0 se debe de escoger muy alta de manera que al principio haya mucho movimiento en el espacio de soluciones. Generalmente el criterio de paro es fijado como un número total de iteraciones en lugar de una temperatura final muy pequeña.

Decisiones específicas

En muchas aplicaciones es muy fácil crear una solución inicial. Luego, esta solución inicial se puede usar como un parámetro para producir soluciones diferentes.

En algunos problemas es posible obtener resultados teóricos relacionados a la estructura del problema; luego una solución vecina se puede generar muy fácilmente. La experiencia ha demostrado que estos casos RS se puede hacer muy eficiente.

Aplicaciones

En Valqui Vidal (1993) y Valqui Vidal (1996) se puede ver una cantidad muy variada de problemas prácticos que se han resuelto usando RS. Este es el caso de problemas de diseño de redes, de localización de plantas, de agrupamientos, de planeamiento de la producción, de estimación estadística, etc.

También se han resuelto problemas combinatorios clásicos. Este es el caso de

- el problema del agente viajero.
- el problema de coloreado de grafos.
- el problema de sucesión de tareas.
- el problema de asignación cuadrática.
- el problema de partición de grafos.
- el problema de la mochila.
- etc.

Además, muy buenos resultados se han obtenido combinando RS con otros métodos heurísticos como es el caso del método de búsqueda tabú (inglés: Tabu search) o el del algoritmo genético (inglés: Genetic algorithm).

Conclusiones

En esta nota se ha presentado una introducción al método heurístico de optimización conocido como RS. Una presentación más profunda se puede encontrar en Pirlot y Valqui Vidal (1996).

Algunas de las ventajas del RS son:

- el proceso de implementación es bastante rápido.
- aplicaciones muy variadas.
- una flexibilidad muy grande.
- produce soluciones de buena calidad.
- es un método transparente.
- requiere mucha experimentación con la computadora.
- fácil de combinar con otros métodos.

La desventaja principal es que requiere mucho tiempo de cálculo en la computadora.

BIBLIOGRAFÍA

1. Kirkpatrick, S. *et al* (1983). *Optimization by Simulated Annealing*, *Science*, 220, 671-680.
2. Černý, V. (1985). *A thermodynamical approach to the travelling salesman problem: an efficient simulation algorithm*, *Journal of Optimization Theory and Applications*, 45, 41-55.
3. Collins, N. E. *et al* (1988). *Simulated annealing - an annotated bibliography*, *AJMMS*, 8, 209-507.

4. Koulamas, C. et al (1994). *A survey of simulated annealing applications to operations research problems*, Omega, 22, 41-56.
5. Azencott, R. (ed.) (1992). *Simulated annealing. Parallelization techniques*. Wiley, NY.
6. Valqui Vidal, R.V. (ed.) (1993). *Applied Simulated Annealing, Lecture Notes in Economics and Mathematical Systems*, Springer, Berlin.
7. Aarts, E., and Korst, J. (1989). *Simulated Annealing and Boltzman Machines: a stochastic approach to combinatorial optimization and neural computing*, Wiley, NY.
8. Valqui Vidal, R. V. (1996). *Simulated annealing applied to combinatorial optimization problems. Special issue of Control and Cybernetics (to appear)*.
9. Pirlot, M., and Valqui Vidal, R. V. (1996). *Simulated Annealing: A tutorial. In [8]*.

FE DE ERRATA (Volumen 1, número 1, junio 1995).

¿Existe la longitud de una varilla?

Pag. 138: Las leyendas de las figuras están intercambiadas.

FE DE ERRATA (Volumen 1, número 1, junio 1995).

El Problema de los Cuatro Colores

La página 131 debe quedar como sigue:

Por otra parte, si p_k es el número de países que poseen k vértices (y, desde luego, también k fronteras), entonces

$$s_{vf} = \sum k p_k, \text{ con } k = 2, \dots, p$$

es la suma de todos los vértices (y la suma de todas las fronteras) que poseen todos los países del mapa en consideración. Ahora, teniendo presente que las fronteras son contadas dos veces (pues cada frontera pertenece a dos países) y que los vértices han sido contados tres veces (pues cada vértice pertenece a tres países de un M_n), podemos escribir,

$$s_{vf} = 2 f = 3 v, \quad p = \sum_{k=2} p_k$$

Reemplazando en la fórmula de Euler

$$\sum_{k=2} p_k + (1/3) \sum_{k=2} k p_k = (1/2) \sum_{k=2} k p_k + 2$$

de donde

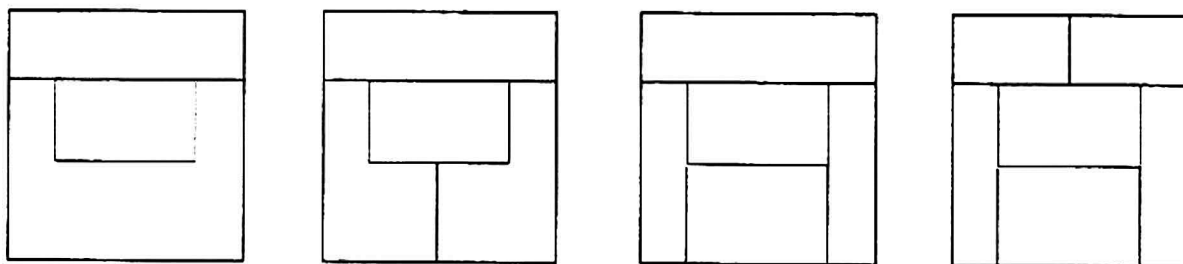
$$\sum_{k=2} (6 - k) p_k = 12$$


y también

$$4 p_2 + 3 p_3 + 2 p_4 + p_5 = 12 + \sum_{k=7} (k - 6) p_k \geq 12$$

donde vemos que es imposible que simultáneamente se anulen p_2 , p_3 , p_4 y p_5 . Es decir, en todo mapa normal -y, por lo tanto, en cada mapa- existen necesariamente países con dos, tres, cuatro o cinco vecinos.

Fig. 6 Regiones donde aparecen países con dos, tres, cuatro o cinco vecinos.



 edición, diagramación e impresión realizada por
ditorial HOZLO S.R.L.

Psje. Santa Rosa 191-501
Telefax: 428-4071

Esta edición consta de 1000 ejemplares

Lima, julio de 1996

