

REVCIUNI

VOLUMEN 7

NÚMERO 2

SETIEMBRE 2003



Revista de la Facultad de Ciencias de la UNI

UNIVERSIDAD NACIONAL DE INGENIERÍA

LIMA - PERÚ

Revista de la Facultad de Ciencias de la UNI - REVCUNI

Publicada por el Instituto de Investigación de la Facultad de Ciencias
de la Universidad Nacional de Ingeniería

Rector: **Ing. Roberto Morales Morales**

Primer Vicerrector: **Ing. Aurelio Padilla Ríos**

Decano de la Facultad de Ciencias: **Dr. Humberto Asmat Azahuanche**

Director del Instituto de Investigación: **Dr. Abel Gutarra Espinoza**

Editor: **Dr. Rosendo Ochoa Jiménez**

Comité Revisor:

Holger Valqui

Orlando Pereyra

Percy Fernández

Leopoldo Paredes

Susana Petrick

Facultad de Ciencias - UNI
Av. Túpac Amaru 210 - Rímac
Teléfono: (0051) (1) 381-3868
E-mail: rochoaj@uni.edu.pe

Ley del Depósito Legal N° 15010199-3597

Revista de la Facultad de
CIENCIAS de la UNI
REVCUNI

Agosto 1960 – Agosto 2003
43 años de la creación de la Facultad de Ciencias

Presentación

Desde el año 1995 el Instituto de Investigación de la Facultad de Ciencias de la Universidad Nacional de Ingeniería ha publicado 11 ejemplares de la revista REVCIUNI. En ellos han colaborado varios de nuestros docentes y alumnos con artículos donde reportan sus investigaciones o difunden temas vinculados a las ciencias básicas y aplicadas.

El número de REVCIUNI que ahora presentamos, tiene algunos cambios formales. Un nuevo diseño en las tapas con mayor variedad de color, así como la inscripción del volumen y el número en el lomo, para una mejor identificación. En cuanto a la temática no han habido modificaciones, la revista acoge trabajos de investigación concluidos o en ejecución, tesis de antegrado y postgrado y artículos de divulgación especializada. En este número, se incluyen temas de física aplicada y de matemáticas.

La publicación de la Revista de la Facultad de Ciencias de la UNI tiene como objetivo principal la difusión de las investigaciones que se realizan en nuestra facultad. Tales investigaciones están reflejados en trabajos relacionados con las tesis licenciatura y trabajos de maestría, así como también resultados importantes de investigación científica de los profesores de la facultad. En la revista se incluye también trabajos de divulgación científica.

En este número se destacan los resultados importantes de las investigaciones del profesor Holger Valqui relacionados con la paradoja de Einstein-Podolsky-Rosen. El trabajo de Juan Carlos Rojas que fue preparado en colaboración con el Instituto de Balseiro de Argentina. También sobresalen los trabajos de proyecto de tesis de Wilfredo Caldas, Oscar Miculicich y el trabajo de Joseph Wilmer relacionado con su tesis de licenciatura. Se incluye también el excelente trabajo del profesor Renato Benazic sobre la teoría de funciones de varias variables complejas.

Sabemos que el hábito de la escritura es inherente al investigador y la principal herramienta de difusión institucional, por ello el Instituto de Investigación considera prioritaria la publicación de la revista de la Facultad y respalda su edición con la frecuencia que el editor considere necesaria.

En el esfuerzo por mejorar cada edición de REVCIUNI contamos con su colaboración, la fortaleza de una institución está en el nivel de participación de quienes la conforman, por ello estamos dispuestos a recibir sus aportes y sugerencias.

Dr. Abel Gutarra Espinoza

Director del Instituto de Investigación

La ecuación de Dirac y la invariancia relativista

Wilfredo Caldas - Orlando Pereyra*

RESUMEN

Una descripción correcta de la mecánica cuántica debe de cumplir con los principios de la relatividad especial, es decir, debe de ser covariante. Luego de una breve introducción histórica sobre la formulación final de la ecuación relativista del electrón, denominada ecuación de Dirac, resolvemos dicha ecuación para el sistema en el cual el electrón está en reposo y usando la invarianza relativista hallamos la solución para un sistema de *referencia* inercial arbitrario. Para este fin construiremos previamente el operador que transforma las funciones de onda de un sistema a otro.

ABSTRACT

A correct description of the quantum mechanics should be complete with the principles of the special relativity, that is to say, it should be covariant. After a brief historical introduction on the final formulation of the relativistic equation of the electron, called Dirac's equation, we solve this equation for the rest frame and using the covariance we find the solution for a moving frame relative to the rest frame. For this purpose we will construct, previously, the operator that transform the wave functions of a one frame to another.

*Grupo de Física Teórica, Facultad de Ciencias - Universidad Nacional de Ingeniería.

1. Introducción

Los principios de la mecánica cuántica y la relatividad especial sirven de base para la construcción de una ecuación de onda covariante, es decir una ecuación de onda invariante ante transformaciones de Lorentz, dicha ecuación como hizo notar Landau [1], no sería una simple generalización de la ecuación de Schrödinger.

En general, el concepto de función de onda no tiene sentido en mecánica cuántica relativista y sólo es posible hablar de ella en el caso de una partícula libre, donde la cantidad de movimiento \vec{p} es constante, dicha afirmación está relacionada con la imposibilidad de realizar mediciones de la cantidad de movimiento en un intervalo de tiempo arbitrariamente corto, debido a que la máxima velocidad de medición es la velocidad de la luz, lo cual haría imposible que el tiempo de medición sea arbitrariamente corto, hecho esencial si se quiere construir una función de onda $\psi(p)$ portadora de toda la información del sistema cuántico. Esta imposibilidad de considerar a la ecuación de Dirac como una ecuación de onda fue establecida en [2] debido a que la ecuación de Dirac no es una ecuación de probabilidad sino una ecuación de campo. La aparición de estados de energía negativa resultan como consecuencia de considerar sino una ecuación de campo. La aparición de estados de energía negativa resultan como consecuencia de considerar a ψ como una función de onda y no como un campo cuántico donde los ψ_μ son operadores de creación y aniquilación y no simples funciones.

En mecánica cuántica, los estados de los sistemas están representados por vectores normalizados $|\psi\rangle$ de un espacio de Hilbert, donde $\langle\varphi|\psi\rangle^2$ es la probabilidad de encontrar el sistema en el estado $|\varphi\rangle$. Los observables físicos son asociados a operadores auto-adjuntos $A=A^\dagger$ que actúan sobre el espacio de Hilbert. El valor esperado del observable A cuando el sistema está en el estado $|\psi\rangle$, es decir, el valor medio para muchas medidas de estados preparados idénticamente es: $\langle\psi|A|\psi\rangle$. La evolución temporal del sistema está gobernada por la ecuación de Schrödinger:

$$i\hbar \frac{\partial}{\partial t} |\psi(t)\rangle = H |\psi(t)\rangle, \quad (1)$$

o equivalentemente:

$$|\psi(t_2)\rangle = U(t_2, t_1) |\psi(t_1)\rangle. \quad (2)$$

Donde H es auto-adjunto, U es unitaria y satisface:

$$i\hbar \frac{\partial}{\partial t_2} |U(t_2, t_1)\rangle = H(t_2) |U(t_2, t_1)\rangle. \quad (3)$$

Por otro lado, la Relatividad Especial establece que las leyes de la naturaleza son independientes de los observadores, si los observadores son de una clase especial de sistemas (sistemas de Galileo), relacionados entre ellos por medio de una transformación del grupo de Poincaré. Este grupo es generado por translaciones espaciales y temporales, rotaciones espaciales, transformaciones de Lorentz o rotaciones espacio-temporales (boost), los cuales relacionan sistemas de referencia moviéndose con velocidad constante. La velocidad de la luz c es absoluta y es la máxima velocidad de cualquier señal. La información originada en (x_0, t_0) llega únicamente a los puntos (x_1, t_1) dentro del cono:

$$c^2(t_1 - t_0)^2 - (x_1 - x_0)^2 \geq 0 \quad : \quad t_1 - t_0 \geq 0. \quad (4)$$

Esta última ecuación es la expresión relativista de causalidad.

Para combinar la invariancia relativista con la mecánica cuántica utilizaremos el principio de correspondencia. En la representación usual del espacio de configuraciones, asociamos los operadores $i\hbar \frac{\partial}{\partial t}$ y $\frac{\hbar}{i} \frac{\partial}{\partial x^i}$ con la energía E y el momentum p^i respectivamente.

Para una partícula masiva libre, la energía está dada en términos del momentum mediante:

$$E^2 = \vec{p}^2 c^2 + m^2 c^4. \quad (5)$$

En adelante utilizaremos el sistema denominado “natural” en el cual $\hbar = c = 1$. La ecuación análoga a la ecuación de Schrödinger correspondiente a la energía relativista (5) se denomina ecuación de Klein-Gordon [3]:

$$\left(\frac{\partial^2}{\partial t^2} - \nabla^2 + m^2 \right) \psi(\vec{x}, t) = 0. \quad (6)$$

Si queremos interpretar a ψ como una función de onda, encontramos que la probabilidad no es necesariamente positiva. La ecuación de continuidad implica:

$$\frac{\partial}{\partial t} \rho + \nabla \cdot \vec{J} \equiv \partial_\mu j^\mu = 0. \quad (7)$$

donde hemos definido el cuadrivector $j^\mu \equiv (j^0 = \rho, j^i)$ como:

$$\begin{aligned}\rho &= \frac{i}{2m} (\psi^* \frac{\partial \psi}{\partial t} - \frac{\partial \psi^*}{\partial t} \psi) \\ \vec{j} &= \frac{i}{2mi} (\psi^* \nabla \psi - (\nabla \psi)^* \psi).\end{aligned}\tag{8}$$

es decir, no podemos asegurar que $\rho > 0$.

Un segundo problema es la existencia de soluciones en ondas planas, las cuales tienen la forma:

$$\psi(\vec{x}, t) = A \exp(i\vec{p} \cdot \vec{x} - iEt).\tag{9}$$

donde $E^2 = \vec{p}^2 + m^2$, es decir, aparecen soluciones con energía negativa $E = -\sqrt{\vec{p}^2 + m^2}$ además de las soluciones con energía positiva $E = \sqrt{\vec{p}^2 + m^2}$.

2. La Ecuación de Dirac

Dado que la ecuación de Klein-Gordon es insatisfactoria físicamente, trataremos de construir una ecuación de onda de la forma:

$$i \frac{\partial \psi}{\partial t} = (\vec{\alpha} \cdot \vec{p} + \beta m) \psi \equiv H\psi,\tag{10}$$

donde ψ es un vector función de onda y $\vec{\alpha}$ y β son matrices hermíticas de tal manera que H resulte hermítica también, lo cual garantizaría la existencia de densidad de probabilidad positiva [3]¹. Para la construcción de esta ecuación tendremos en cuenta estos puntos:

1. La ecuación debe de ser covariante.
2. Las componentes de ψ deben satisfacer la ecuación de Klein-Gordon, así que la onda plana con $E^2 = \vec{p}^2 + m^2$ debe de ser solución.

¹ Lo que garantiza, en realidad, el signo tanto de la carga como de la energía es la manera como se transforman los campos, es decir, si estos son espinores o tensores. La probabilidad es definida positiva si el campo es espinorial pero la energía no, mientras que si el campo es tensorial entonces la energía es definida positiva pero la probabilidad no [4].

Dirac [5], encontró que las matrices α_i y β deben satisfacer las siguientes relaciones.

$$\{\alpha_i, \alpha_j\} = 0, \quad i \neq j, \quad \{\alpha_i, \beta\} = 0, \quad \alpha_i^2 = \beta^2 = 1. \quad (11)$$

donde $\{A, B\} \equiv AB + BA$ se le denomina anticonmutador de A y B .

Es fácil verificar que con las condiciones anteriores (10) es equivalente a la ecuación de Klein-Gordon

$$-\frac{\partial^2}{\partial t^2} \psi = (\vec{\alpha} \cdot \vec{p} + \beta m)^2 \psi = (-\nabla^2 + m^2) \psi. \quad (12)$$

Introduzcamos ahora la notación γ^μ :

$$\gamma^0 = \beta, \quad \gamma^i = \beta \alpha^i \quad i = 1, 2, 3. \quad (13)$$

Entonces las matrices γ^μ satisfacen la relación:

$$\{\gamma^\mu, \gamma^\nu\} = 2g^{\mu\nu}, \quad (14)$$

donde, $g^{\mu\nu} = [1 \quad -1 \quad -1 \quad -1]$, es la métrica de Minkowsky.

Si utilizamos estas definiciones y reescribimos la ecuación (10), obtenemos lo que llamaremos Ecuación de Dirac:

$$i\gamma^\mu \frac{\partial \psi(x)}{\partial x^\mu} - m\psi(x) = 0, \quad \text{donde,} \quad x \equiv (t, \vec{x}) \quad (15)$$

Las matrices γ^μ se denominan matrices de Dirac; la menor dimensión [6] de estas matrices es cuatro, y pueden escribirse en función de las matrices 2x2 de Pauli:

$$\gamma^0 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \quad \gamma^i = \begin{pmatrix} 0 & \sigma^i \\ -\sigma^i & 0 \end{pmatrix} \quad (16)$$

También existen otras formas de representar las matrices de Dirac de manera tal que aún cumplan con (14), i.e. la representación quirial:

$$\gamma^0 = \begin{pmatrix} 0 & -1 \\ -1 & 0 \end{pmatrix} \quad \gamma^i = \begin{pmatrix} 0 & \sigma^i \\ -\sigma^i & 0 \end{pmatrix} \quad (17)$$

donde $i = 1..3$

3. Invariancia relativista

De acuerdo al principio de relatividad la ecuación de Dirac mantiene su forma en dos sistemas de referencia si éstos están relacionados por una transformación de Lorentz, consideremos Λ dicha transformación, nuestro sistema físico será descrito por una función ψ en el primer sistema y por ψ' en el sistema transformado, ambas satisfacen la Ecuación de Dirac:

$$i\gamma^\mu \frac{\partial \psi(x)}{\partial x^\mu} - m\psi(x) = 0 \quad (18)$$

$$i\gamma^\mu \frac{\partial \psi'(x')}{\partial x'^\mu} - m\psi'(x') = 0 \quad \text{donde} \quad x' = \Lambda x. \quad (19)$$

Existe una relación local entre ψ y ψ' de tal forma que el observador en el segundo sistema de referencia pueda reconstruir ψ' si conoce ψ , asumamos que esta relación es lineal:

$$\psi'(x') = S(\Lambda)\psi(x), \quad (20)$$

$S(\Lambda)$ es una matriz de transformación para la función de onda que depende de la transformación de Lorentz para las coordenadas, reemplazando (20) en (19) obtenemos:

$$i\gamma^\mu \frac{\partial S(\Lambda)\psi(x)}{\partial x'^\mu} - mS(\Lambda)\psi(x) = 0. \quad (21)$$

Los elementos de la matriz de transformación son:

$$\frac{\partial x'^\mu}{\partial x^\nu} = (\Lambda)^\mu_\nu, \quad \frac{\partial x^\mu}{\partial x'^\nu} = (\Lambda^{-1})^\mu_\nu. \quad (22)$$

Escribamos (21) de tal forma que aparezcan los elementos de la transformación inversa

$$i\gamma^\mu (\Lambda)^\nu_\mu \frac{\partial S(\Lambda)\psi(x)}{\partial x^\nu} - mS(\Lambda)\psi(x) = 0. \quad (23)$$

Esta ecuación necesita ser consecuente con la ecuación (18) en el sistema de referencia inicial, entonces debe satisfacerse la relación:

$$S(\Lambda)\gamma^\mu S^{-1}(\Lambda) = (\Lambda^{-1})^\mu_\nu \gamma^\nu. \quad (24)$$

A continuación vamos a encontrar la expresión explícita del operador $S(\Lambda)$ para esto multiplicaremos $S^{-1}(\Lambda)$ por la izquierda en la ecuación (23)

$$iS^{-1}(\Lambda)\gamma^\mu(\Lambda)_\mu^\nu S(\Lambda)\frac{\partial\psi(x)}{\partial x^\nu} - m\psi(x) = 0.$$

como $S(\Lambda)$ no depende explícitamente de las coordenadas podemos extraerlo de la derivada parcial, obteniendo:

$$\gamma^\mu = S^{-1}(\Lambda)\gamma^\nu(\Lambda)_\nu^\mu S(\Lambda).$$

o lo que es equivalente:

$$S(\Lambda)\gamma^\nu S^{-1}(\Lambda) = \gamma^\mu(\Lambda)_\mu^\nu. \quad (25)$$

Mediante esta relación construiremos $S(\Lambda)$ utilizando para ello las propiedades infinitesimales de Λ , las cuales pueden ser escritas como:

$$(\Lambda)_\nu^\mu = g_\nu^\mu + \delta\omega_\nu^\mu, \quad (\Lambda^{-1})_\nu^\mu = g_\nu^\mu - \delta\omega_\nu^\mu. \quad (26)$$

$\delta\omega_{\mu\nu}$ es un tensor antisimétrico infinitesimal ($\omega_{\mu\nu} = -\omega_{\nu\mu}$) [3][8].

Escribamos $S(\Lambda)$ en serie:

$$S(\Lambda) = S_0 - iS_1 + \dots, \quad S^{-1}(\Lambda) = S_0 + iS_1 + \dots \quad (27)$$

El primer término S_0 en (27) debe ser la identidad, que correspondería a la transformación $\Lambda = g$. El segundo término, S_1 es la primera aproximación, que en principio debería ser función lineal $\omega_{\mu\nu}$ por tanto podemos escribir en primer orden:

$$S(\Lambda) = 1 + iS_1, \quad S^{-1}(\Lambda) = 1 - iS_1. \quad (28)$$

Remplazando estas expresiones en la ecuación (25) y considerando sólo términos de primer orden

$$(1 + iS_1)\gamma^\alpha(1 - iS_1) = \gamma^\nu(g_\nu^\alpha + \delta\omega_\nu^\alpha).$$

$$Y_i[S_1, \gamma^\alpha] = \lambda^\nu \delta\omega_\nu^\alpha. \quad (29)$$

Que también puede escribirse como:

$$i[S_1, \gamma^\alpha] = g^{\alpha\mu} \gamma^\nu \delta\omega_{\mu\nu}. \quad (30)$$

Ahora haremos que aparezca γ^α en la expresión de la derecha de manera parecida a $[S_1, \gamma^\alpha]$ con el fin de identificar S_1 .

Utilizando el hecho que μ, ν son índices mudos podemos escribir la ecuación anterior como:

$$i[S_1, \gamma^\alpha] = \frac{1}{2} (g^{\alpha\mu} \gamma^\nu \delta\omega_{\mu\nu} + g^{\alpha\nu} \gamma^\mu \delta\omega_{\nu\mu}).$$

pero $\delta\omega_{\mu\nu} = -\delta\omega_{\nu\mu}$

$$i[S_1, \gamma^\alpha] = \frac{1}{2} (g^{\alpha\mu} \gamma^\nu - \gamma^\mu g^{\alpha\nu}) \delta\omega_{\mu\nu}.$$

y como $\frac{1}{2} \{\gamma^\epsilon, \gamma^\mu\} = g^{\alpha\mu}$ entonces:

$$i[S_1, \gamma^\alpha] = \frac{1}{2} \left(\frac{1}{2} \{\gamma^\alpha, \gamma^\mu\} \gamma^\nu - \gamma^\mu \frac{1}{2} \{\gamma^\alpha, \gamma^\nu\} \right) \delta\omega_{\mu\nu}$$

utilizando la propiedad $[AB, C] = A\{B, C\} - \{A, C\} B$ resulta:

$$i[S_1, \gamma^\alpha] = -\frac{1}{4} [\gamma^\mu \gamma^\nu, \gamma^\alpha] \delta\omega_{\mu\nu}$$

Entonces:

$$[S_1, \gamma^\alpha] = \frac{i}{4} [\delta\omega_{\mu\nu} \gamma^\mu \gamma^\nu, \gamma^\alpha]$$

Comparando podemos ver que:

$$\begin{aligned} S_1 &= \frac{i}{4} \delta\omega_{\mu\nu} \gamma^\mu \gamma^\nu = \frac{i}{8} (\delta\omega_{\mu\nu} \gamma^\mu \gamma^\nu + \delta\omega_{\nu\mu} \gamma^\nu \gamma^\mu) = \frac{i}{8} (\gamma^\mu \gamma^\nu - \gamma^\nu \gamma^\mu) \delta\omega_{\mu\nu}, \\ &= \frac{i}{8} [\gamma^\mu, \gamma^\nu] \delta\omega_{\mu\nu}. \end{aligned} \quad (31)$$

Entonces en primera aproximación el operador de transformación para la función de onda de Dirac es:

$$S(\Lambda) = 1 - \frac{1}{8} [\gamma^\mu, \gamma^\nu] \delta\omega_{\mu\nu}, \quad S^{-1}(\Lambda) = 1 + \frac{1}{8} [\gamma^\mu, \gamma^\nu] \delta\omega_{\mu\nu}. \quad (32)$$

Debido a la forma diferencial de $S(\Lambda)$ en (32), podemos escribir la siguiente expresión para cantidades $\omega_{\mu\nu}$ no infinitesimales:

$$S(\Lambda) = \exp\left(-\frac{1}{8}[\gamma^\mu, \gamma^\nu] \omega_{\mu\nu}\right) = \exp\left(\frac{i}{4}\sigma^{\mu\nu} \omega_{\mu\nu}\right). \quad (33)$$

donde hemos definido:

$$\sigma^{\mu\nu} \equiv \frac{i}{2}[\gamma^\mu, \gamma^\nu]. \quad (34)$$

4. Forma explícita del operador $S(\Lambda)$

Sea Λ una transformación de Lorentz x - t representado por la matriz:

$$\Lambda_\nu^\mu = \begin{pmatrix} \cosh \varepsilon & \sinh \varepsilon & 0 & 0 \\ \sinh \varepsilon & \cosh \varepsilon & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \quad \text{donde} \quad \tanh \varepsilon = v. \quad (35)$$

La última expresión podemos escribirla en forma diferencial:

$$\Lambda_\nu^\mu = g_\nu^\mu + \delta\omega_\nu^\mu, \quad (36)$$

donde en este caso:

$$\delta\omega = \delta\varepsilon \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}. \quad (37)$$

Sea ahora Λ una rotación espacial, por ejemplo alrededor del eje z , entonces estará representado por la matriz:

$$\Lambda_\nu^\mu = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos \theta & -\sin \theta & 0 \\ 0 & \sin \theta & \cos \theta & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \quad (38)$$

que en forma diferencial se escribe como:

$$\Lambda_\nu^\mu = g_\nu^\mu + \delta\omega_\nu^\mu, \quad (39)$$

donde en este caso:

$$\delta\omega_{\nu}^{\mu} = \delta\theta \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}. \quad (40)$$

En general, para una transformación arbitraria del grupo de Lorentz, los parámetros de transformación, forman una matriz antisimétrica:

$$\delta\omega_{\mu\nu} = g_{\mu\alpha} \delta\omega_{\nu}^{\alpha} \quad (41)$$

Los parámetros correspondientes a la velocidad $|\vec{v}| = \tanh|\vec{\epsilon}|$ están relacionados con los términos $\omega_{0i} = \epsilon_i$ mientras que los ángulos de rotación alrededor de los ejes x, y, z están determinados por $\theta^k = \frac{1}{2} \epsilon^{\nu k} \omega_{\nu}$.

Para hallar una forma explícita del operador $S(\Lambda)$ utilizaremos la representación quiral de las matrices de Dirac, las cuales satisfacen las siguientes relaciones [3]:

$$\sigma^{0i} = \frac{i}{2} [\gamma^0, \gamma^i] = -i\alpha^i = -i \begin{pmatrix} \sigma_i & 0 \\ 0 & -\sigma_i \end{pmatrix} \quad (42)$$

$$\sigma^{\nu} = \frac{i}{2} [\gamma^i, \gamma^j] = \epsilon^{\nu k} \begin{pmatrix} \sigma_k & 0 \\ 0 & \sigma_k \end{pmatrix} \quad (43)$$

donde σ_i son las matrices de Pauli y $\epsilon_{\nu k}$ es el tensor antisimétrico de Levi-Cevita. Descomponiendo (33):

$$S(\Lambda) = \exp\left(\frac{i}{4} [\sigma^{00} \omega_{00} + \sigma^{0i} \omega_{0i} + \sigma^{\nu} \omega_{\nu}]\right) \quad (44)$$

Debido a la antisimetría de $\omega_{\mu\nu}$ y de acuerdo a (42) y (43) obtenemos:

$$S(\Lambda) = \exp\left(\frac{i}{4} [-2i\alpha^i \omega_{0i} + \epsilon^{\nu k} \begin{pmatrix} \sigma_k & 0 \\ 0 & \sigma_k \end{pmatrix} \omega_{\nu}]\right) \quad (45)$$

si definimos $\vec{\alpha} \equiv (\alpha_1, \alpha_2, \alpha_3)$, $\vec{\epsilon} \equiv (\omega_{01}, \omega_{02}, \omega_{03})$, $\vec{\theta} \equiv (\omega_{23}, \omega_{31}, \omega_{12})$ y $\vec{\sigma} \equiv (\sigma_1, \sigma_2, \sigma_3)$ la expresión anterior se reduce a:

$$S(\Lambda) = \exp\left(\frac{1}{2} \vec{\alpha} \cdot \vec{\epsilon} + \frac{i}{2} \vec{\theta} \cdot \vec{\sigma}\right) \quad (46)$$

El primer término corresponde a las transformaciones espacio temporales (boost) y el segundo término es la forma usual para la transformación de espinores en la mecánica cuántica no relativista,

relacionada con rotaciones puramente espaciales; las cantidades $\vec{\varepsilon}$ son los parámetros de los boost relacionados con la velocidad física \vec{v} por $\tanh \varepsilon = v$, donde $v = |\vec{v}|$. Mientras que $\theta = |\vec{\theta}|$ es el ángulo de giro tridimensional con respecto al eje $\vec{n} = \frac{\vec{\theta}}{\theta}$.

El primer término de la ecuación (46) puede ser expandido en serie

$$S(\Lambda) = 1 + \frac{1}{2} \vec{\alpha} \cdot \vec{\varepsilon} + \frac{1}{2!} \left(\frac{1}{2} \vec{\alpha} \cdot \vec{\varepsilon} \right)^2 + \frac{1}{3!} \left(\frac{1}{2} \vec{\alpha} \cdot \vec{\varepsilon} \right)^3 + \dots = \cosh \frac{\varepsilon}{2} + (\vec{\alpha} \cdot \vec{n}) \sinh \frac{\varepsilon}{2} \quad (47)$$

donde hemos utilizado la propiedad $\alpha_i^2 = 1$ y definido $\vec{n} \equiv \frac{\vec{\varepsilon}}{\varepsilon}$, el cual es un vector unitario que tiene la misma dirección del vector de propagación \vec{p} .

El segundo término en (46) sólo interviene si se realizan rotaciones espaciales sobre el sistema de coordenadas. No las usaremos aquí, en tanto que sólo nos interesa realizar sobre el sistema de coordenadas, una transformación de Lorentz.

5. Solución de la ecuación de Dirac

La ecuación de Dirac, por ser una ecuación que describe una partícula libre, admite como solución ondas planas de la forma:

$$\Psi^+(x) = e^{-ipx} u(p), \quad \Psi^-(x) = e^{ipx} v(p), \quad p_0 > 0$$

La ecuación de Dirac implica que:

$$(\gamma^\nu p_\nu - m)u(p) = 0, \quad (\gamma^\nu p_\nu + m)v(p) = 0$$

Si asumimos que la partícula es masiva, $m \neq 0$ entonces en el sistema de referencia propio, $p^\nu = (m, \vec{0})$, la ecuación anterior se reduce a:

$$(\gamma^0 - 1)u(m, \vec{0}) = 0, \quad (\gamma^0 + 1)v(m, \vec{0}) = 0$$

Existen dos soluciones linealmente independientes tanto para u como para v , en la representación usual de Dirac dichas soluciones pueden escribirse como:

$$u^{(1)}(m, \vec{0}) = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \quad u^{(2)}(m, \vec{0}) = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \quad v^{(1)}(m, \vec{0}) = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \quad v^{(2)}(m, \vec{0}) = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}. \quad (48)$$

Estas son las soluciones de la ecuación de Dirac, en la representación de momentum, en el sistema de referencia propio. Ahora utilizaremos la exposición (46) para hallar la solución en un sistema de referencia inercial que tiene una velocidad \bar{v} con respecto del sistema propio.

El operador que transforma la función de onda de un sistema a otro está dado por (47). Entonces buscaremos una expresión explícita de este operador:

$$S(\Lambda) = \cosh \frac{\omega}{2} + (\vec{\alpha} \cdot \vec{n}) \sinh \frac{\omega}{2} \quad (49)$$

$$\text{donde } \vec{n} = \frac{\vec{\omega}}{\omega} \quad \tan \omega = v$$

de la relación anterior se puede mostrar que:

$$\sinh \omega = \frac{v}{\sqrt{1-v^2}}, \quad \cosh \omega = \frac{1}{\sqrt{1-v^2}},$$

Utilizando las identidades:

$$\sinh \frac{\omega}{2} = \frac{1}{\sqrt{2}} \sqrt{\cosh \omega - 1}, \quad \cosh \frac{\omega}{2} = \frac{1}{\sqrt{2}} \sqrt{\cosh \omega + 1},$$

y las relaciones $E^2 = p^2 + m^2$ y $v = \frac{|\vec{p}|}{E}$ deducimos las siguientes igualdades:

$$\sinh \frac{\omega}{2} = \sqrt{\frac{E-m}{2m}}, \quad \cosh \frac{\omega}{2} = \sqrt{\frac{E+m}{2m}}. \quad (50)$$

Entonces el operador de transformación adquiere la forma:

$$S(\Lambda) = \sqrt{\frac{E+m}{2m}} + (\vec{\alpha} \cdot \vec{n}) \sqrt{\frac{E-m}{2m}}, \quad (51)$$

$$\text{con } \vec{n} = \frac{\vec{p}}{p}, \quad |\vec{p}| = \sqrt{E^2 - m^2}.$$

Utilizando la representación de Dirac para $\vec{\alpha}$ llegamos a la siguiente fórmula:

$$S(\Lambda) = \sqrt{\frac{E+m}{2m}} \begin{pmatrix} 1 & \frac{\vec{\sigma} \cdot \vec{p}}{E+m} \\ \frac{\vec{\sigma} \cdot \vec{p}}{E+m} & 1 \end{pmatrix}.$$

Esta última ecuación podemos escribirla en forma explícita:

$$S(\Lambda) = \sqrt{\frac{E+m}{2m}} \begin{pmatrix} 1 & 0 & \frac{p_z}{E+m} & \frac{p_x - ip_y}{E+m} \\ 0 & 1 & \frac{p_x + ip_y}{E+m} & \frac{-p_z}{E+m} \\ \frac{p_z}{E+m} & \frac{p_x - ip_y}{E+m} & 1 & 0 \\ \frac{p_x + ip_y}{E+m} & \frac{-p_z}{E+m} & 0 & 1 \end{pmatrix} \quad (53)$$

Esta matriz transforma las soluciones de la ecuación de Dirac en el sistema de referencia propio (48) en soluciones desde un sistema de referencia inercial con velocidad v , es decir transforma la función de onda de (48) en:

$$u^{(1)}(m, \vec{p}) = \sqrt{\frac{E+m}{2m}} \begin{pmatrix} 1 \\ 0 \\ \frac{p_z}{E+m} \\ \frac{p_x + ip_y}{E+m} \end{pmatrix}, \quad u^{(2)}(m, \vec{p}) = \sqrt{\frac{E+m}{2m}} \begin{pmatrix} 0 \\ 1 \\ \frac{p_x - ip_y}{E+m} \\ \frac{-p_z}{E+m} \end{pmatrix} \quad (54)$$

$$v^{(1)}(m, \vec{p}) = \sqrt{\frac{E+m}{2m}} \begin{pmatrix} \frac{p_z}{E+m} \\ \frac{p_x + ip_y}{E+m} \\ 1 \\ 0 \end{pmatrix}, \quad v^{(2)}(m, \vec{p}) = \sqrt{\frac{E+m}{2m}} \begin{pmatrix} \frac{p_x - ip_y}{E+m} \\ \frac{-p_x}{E+m} \\ 0 \\ 1 \end{pmatrix} \quad (55)$$

Estos vectores de onda están normalizados como es fácil comprobar si se calcula:

$$\bar{u}^1(p)u^1(p) = u^+(p) \gamma^0 u(p) = 1 \quad (56)$$

Conclusión

Se usó la invariancia relativista de manera didáctica para hallar el operador de transformación (46) que asocia a cada observador un estado, el cual se encuentra completamente determinado, si se conoce el momentum y energía en cierto instante, lo cual es característico de una partícula libre.

En mecánica cuántica no relativista la función de onda asociada a un electrón es un espinor de dos componentes, cuya ley de transformación está dada por:

$\Psi' = R\Psi$ donde:

$$R = \cos \frac{\omega}{2} + (\vec{\sigma} \cdot \vec{n}) \operatorname{isen} \frac{\omega}{2} \quad (57)$$

El carácter espinorial de la ecuación de Dirac se manifiesta claramente en la forma de la matriz de transformación $S(\Lambda)$, que transforma espinores de cuatro componentes llamados bi-espinores:

$\Psi' = S(\Lambda)\Psi$ donde:

$$S(\Lambda) = \cosh \frac{\omega}{2} + (\vec{\alpha} \cdot \vec{n}) \sinh \frac{\omega}{2} \quad (58)$$

La aparición de funciones elípticas en (58) en lugar de funciones circulares se debe a que mientras los espinores de la mecánica cuántica no relativista se transforma sólo bajo rotaciones espaciales, las funciones de onda de la mecánica cuántica relativista se transforman, además, bajo rotaciones espacio-temporales (Boost), como lo demuestra manifiestamente (46).

Referencias

1. *L.D. Landau and E.M. Lifshitz*, "Quantum Mechanics" Mir, Moscu (1980).
2. *W. Pauli and V. Weisskopf*, *Helv. Phys. Acta* 7, 709-731 (1934).
3. *C. Itzykson*, "Quantum Field Theory" McGraw-Hill (1980).
4. *W. Pauli*, *Phys. Rev.* 58, 630 (1940).
5. *P.A.M. Dirac*, *Proc. R. Soc A* 118, 654 (1928).
6. *P.A.M. Dirac*, *Proc. R. Soc A* 112, 661 (1926).
7. *J.L. Lopes*, "Introducción a la Electrodinámica Cuántica" Trillas, México (1977).
8. *J.D. Borken and S.D. Drell* "Relativistic Quantum Mechanics" McGraw-Hill, New York (1964).

Funciones de Estado y Polarización de Fotones

Oscar Miculicich, H. G. Valqui*

RESUMEN

Se utiliza la transformada de Fourier para obtener los estados de polarización del fotón a partir de los campos $E(\vec{r},t)$ y $H(\vec{r},t)$ que son soluciones de las ecuaciones de Maxwell. Primero se representan las ecuaciones de Maxwell en el espacio transformado, cuyas nuevas soluciones $\vec{E}(k,t)$ y $\vec{H}(k,t)$ son funciones que satisfacen las ecuaciones de Maxwell modificadas (EMM). Al transformar la expresión clásica de la energía del campo electromagnético, se obtiene otra expresión que, por su forma, es identificada con el valor esperado de la energía del fotón, tomando como funciones de estado ciertas soluciones de las EMM. Esta elección de las *funciones de estado* es reforzada al mostrar la coincidencia entre el valor esperado del momento lineal del fotón y la correspondiente expresión clásica del campo electromagnético; el refuerzo es aún mayor en el caso del momento angular, asociado al espín del fotón.

ABSTRACT

Starting from the electromagnetic fields, \vec{E} and \vec{H} , solution of the Maxwell equations, one applies the Fourier Transform to get the photon's polarization states. First, we write the transformed Maxwell equations, TME, and its solutions \vec{E} and \vec{H} . The new formula for the electromagnetic's energy density is interpreted as the expected value of the photon's energy, where the state function is a solution of TME.

Such choice of the state function shows its advantage in the construction of the expected value corresponding to the photon's linear momentum, and more so in the construction of the expected value of the photon's angular momentum, where, as a bonus, appears the photon's spin.

* Facultad de Ciencias - Universidad Nacional de Ingeniería.

1. Representación en el espacio momentum

1.1 Ecuaciones de Maxwell

Se asume que las soluciones de las ecuaciones de Maxwell en el vacío consisten de dos funciones vectoriales del vector posición y del tiempo: $\vec{E}(\vec{r}, t)$ y $\vec{H}(\vec{r}, t)$. Estas funciones pueden ser escritas como una superposición de ondas planas:

$$\vec{E}(\vec{r}, t) = \frac{1}{(2\pi)^{3/2}} \int dk^3 \vec{\mathcal{E}}(\vec{k}, t) e^{i\vec{k} \cdot \vec{r}} \quad (1)$$

$$\vec{H}(\vec{r}, t) = \frac{1}{(2\pi)^{3/2}} \int dk^3 \vec{\mathcal{H}}(\vec{k}, t) e^{i\vec{k} \cdot \vec{r}}. \quad (2)$$

$\vec{E}(\vec{r}, t)$ y $\vec{H}(\vec{r}, t)$ obedecen las ecuaciones de Maxwell, ahora se determinará que ecuaciones cumplen $\vec{\mathcal{E}}(\vec{k}, t)$ y $\vec{\mathcal{H}}(\vec{k}, t)$. Puesto que el momentum de un fotón está dado por $\vec{p} = \hbar \vec{k}$, la representación usando coordenadas k_1 , k_2 y k_3 se denominará representación en el *K-espacio o espacio momentum*.

En el espacio libre tanto la densidad de carga ρ como el vector densidad de corriente \vec{J} son nulas; en este caso las ecuaciones de Maxwell toman la siguiente forma:

$$\nabla \times \vec{E} = -\mu \partial_t \vec{H} \quad (3)$$

$$\nabla \cdot \vec{E} = 0 \quad (4)$$

$$\nabla \times \vec{H} = \varepsilon \partial_t \vec{E} \quad (5)$$

$$\nabla \cdot \vec{H} = 0 \quad (6)$$

donde $\partial_t \equiv \frac{\partial}{\partial t}$; μ , ε son las constantes magnéticas y dieléctrica del vacío respectivamente.

Al reemplazar (1) y (2) en las ecuaciones anteriores, se obtienen:

$$i\vec{k} \times \vec{\mathcal{E}}(\vec{k}, t) = -\mu \partial_t \vec{\mathcal{H}}(\vec{k}, t) \quad (7)$$

$$\vec{k} \cdot \vec{\mathcal{E}}(\vec{k}, t) = 0 \quad (8)$$

$$i\vec{k} \times \vec{\mathcal{H}}(\vec{k}, t) = \varepsilon \partial_t \vec{\mathcal{E}}(\vec{k}, t) \quad (9)$$

$$\vec{k} \cdot \vec{\mathcal{H}}(\vec{k}, t) = 0 \quad (10)$$

que representan a las ecuaciones de Maxwell en el *K-espacio*.

Ahora, sea $\vec{V}(\vec{r})$ una función vectorial de la posición, cuya transformada de Fourier se puede escribir como:

$$\vec{V}(\vec{r}) = \frac{1}{(2\pi)^{3/2}} \int d\vec{k}^3 \vec{\mathcal{V}}(\vec{k}) e^{i\vec{k} \cdot \vec{r}},$$

donde la exigencia de que $\vec{V}(\vec{r})$ sea real, $\vec{V}^*(\vec{r}) = \vec{V}(\vec{r})$ implica que:

$$\int d\vec{k}^3 \vec{\mathcal{V}}(\vec{k}) e^{i\vec{k} \cdot \vec{r}} = \int d\vec{k}^3 \vec{\mathcal{V}}^*(\vec{k}) e^{-i\vec{k} \cdot \vec{r}},$$

y haciendo un cambio de \vec{k} por $-\vec{k}$ en el lado izquierdo de la expresión anterior, y por razones de independencia lineal, se verifique:

$$\vec{\mathcal{V}}^*(\vec{k}) = \vec{\mathcal{V}}(-\vec{k}) \quad (11)$$

La igualdad anterior es llamada *condición de realidad* de los coeficientes de Fourier.

Puesto que $\vec{E}(\vec{r}, t)$ y $\vec{H}(\vec{r}, t)$ son campos reales; entonces los campos transformados $\vec{\mathcal{E}}(\vec{k}, t)$ y $\vec{\mathcal{H}}(\vec{k}, t)$ además de las ecuaciones (7) a (10), deben satisfacer la condición (11), es decir:

$$\vec{\mathcal{E}}^*(\vec{k}, t) = \vec{\mathcal{E}}(-\vec{k}, t) \quad (12)$$

$$\vec{\mathcal{H}}^*(\vec{k}, t) = \vec{\mathcal{H}}(-\vec{k}, t) \quad (13)$$

1.2 Las variables de campo en el K-espacio

En el espacio real es posible obtener las ecuaciones de onda para los campos $\vec{E}(\vec{r}, t)$ y $\vec{H}(\vec{r}, t)$ a partir de las ecuaciones de Maxwell (3) a (6). Similarmente, en el K-espacio al acoplar las ecuaciones (7) a (10) se obtienen las ecuaciones:

$$(\partial_t^2 + k^2 c^2) \vec{\mathcal{E}}(\vec{k}, t) = \vec{0} \quad (14)$$

$$(\partial_t^2 + k^2 c^2) \vec{\mathcal{H}}(\vec{k}, t) = \vec{0} \quad (15)$$

que representan a las ecuaciones de onda en el espacio momentum.

Las dos ecuaciones anteriores son coincidentes, por tal razón ahora se analiza las soluciones $\vec{f}(\vec{k}, t)$ de la ecuación:

$$(\partial_t^2 + k^2 c^2) \vec{f}(\vec{k}, t) = \vec{0} \quad (16)$$

que puede ser escrita de la siguiente forma:

$$(\partial_t + ikc)(\partial_t - ikc) \vec{f}(\vec{k}, t) = \vec{0}. \quad (17)$$

De (17), se puede ver que la solución general de la ecuación (16) será una combinación lineal de las soluciones independientes de las siguientes ecuaciones:

$$(\partial_t + ikc) \bar{f}(\bar{k}, t) = \bar{0}, \quad (18)$$

$$(\partial_t - ikc) \bar{f}'(\bar{k}, t) = \bar{0}. \quad (19)$$

Ahora, sea $\bar{f}(\bar{k}, t)$ solución de la ecuación (18); luego, si se realiza el cambio de t por $-t$, o se toma el conjugado complejo de la ecuación (18) se verifica que $\bar{f}(\bar{k}, -t)$ y $\bar{f}^*(\bar{k}, t)$ son soluciones de la ecuación (19) (que a lo más difieren en un factor por tratarse de una ecuación diferencial de primer orden). En cambio, si en la ecuación (18) se hace el cambio de \bar{k} por $-\bar{k}$ y puesto que $k = |\bar{k}| = |-\bar{k}|$ se verifica que $\bar{f}(-\bar{k}, t)$ es solución para la misma ecuación. Por lo anterior, si se tiene la solución para una de las ecuaciones (18) ó (19), entonces es posible obtener la solución para la ecuación restante.

Así la solución general de la ecuación (16) puede expresarse como:

$$\bar{\mathcal{E}}(\bar{k}, t) = A\bar{f}(\bar{k}, t) + B\bar{f}^*(-\bar{k}, t) \quad (20)$$

donde A y B son coeficientes independientes del tiempo pero pueden depender de k . La expresión anterior debe cumplir la condición de realidad; es decir: $\bar{\mathcal{E}}^*(\bar{k}, t) = \bar{\mathcal{E}}(-\bar{k}, t)$, de donde se obtiene:

$$\bar{\mathcal{E}}(\bar{k}, t) = A\bar{f}(\bar{k}, t) + A^*\bar{f}^*(-\bar{k}, t) \quad (21)$$

que al derivar parcialmente con respecto al tiempo y reemplazando las ecuaciones (18) y (19) se obtiene:

$$\partial_t \bar{\mathcal{E}}(\bar{k}, t) = -ikc[A\bar{f}(\bar{k}, t) - A^*\bar{f}^*(-\bar{k}, t)] \quad (22)$$

A partir de las ecuaciones (21) y (22) se verifica que:

$$\bar{f}(\bar{k}, t) = \frac{1}{2A} [\bar{\mathcal{E}}(\bar{k}, t) - \frac{1}{ikc} \partial_t \bar{\mathcal{E}}(\bar{k}, t)] \quad (23)$$

y si se multiplica escalarmente esta expresión por \bar{k} y aplicando (8), se verifica que:

$$\bar{k} \cdot \bar{f}(\bar{k}, t) = 0, \quad (24)$$

que es conocida como la *condición de transversabilidad*.

1.3 Las funciones de estado en el K-espacio

Con la ecuación (23) se ha mostrado una relación entre una función $\vec{f}(\vec{k}, t)$ en el espacio momentum y los coeficientes de Fourier $\vec{E}(\vec{k}, t)$; por consiguiente con el campo eléctrico real $\vec{E}(\vec{r}, t)$. Ahora se reconsidera la ecuación (18).

$\partial_t \vec{f}(\vec{k}, t) = -ikc\vec{f}(\vec{k}, t)$ que, al multiplicar por $i\hbar$ ambos miembros de la igualdad, se obtiene:

$$i\hbar \partial_t \vec{f}(\vec{k}, t) = \hbar kc \vec{f}(\vec{k}, t) \quad (25)$$

lo que al comparar con la ecuación de Schrödinger:

$$i\hbar \partial_t |S\rangle = H|S\rangle, \quad (26)$$

hace posible identificar a \vec{f} como una *función de estado* en el espacio momentum si se toma como Hamiltoniano al operador multiplicativo $H = k\hbar c$.

Ahora, se obtendrá una expresión para la constante A en la expresión (23) de tal forma que \vec{f} quede completamente determinado. El *valor esperado de la energía* en el K-espacio está dado por:

$$\langle H \rangle = \langle S | k\hbar c | S \rangle = \int d^3k \vec{f}^*(\vec{k}, t) \cdot (k\hbar c) \vec{f}(\vec{k}, t), \quad (27)$$

lo que debe corresponder a la energía clásica del campo electromagnético:

$$W = \frac{1}{2} \int d^3r [\epsilon \vec{E}^2 + \mu \vec{H}^2]. \quad (28)$$

Al reemplazar las ecuaciones (1) y (2) en la expresión anterior, luego del cálculo, simplificar y usando las ecuaciones (21) y (22) se verifica que:

$$W = \int dk^3 \vec{f}^*(\vec{k}, t) \cdot (2\epsilon |A|^2) \vec{f}(\vec{k}, t), \quad (29)$$

en donde al demandar que $W = \langle H \rangle$, la expresión de A debe satisfacer:

$$|A| = \left(\frac{k\hbar c}{2\epsilon} \right)^{1/2} \quad (30)$$

con lo cual $\vec{f}(\vec{k}, t)$ puede ser considerada como la función de estado del fotón en el espacio momentum.

2. Momento lineal y angular en el K -espacio

2.1 Momento lineal del fotón

En esta sección se verificará que en el espacio momentum las funciones $\vec{f}(\vec{k}, t)$ dadas por la ecuación (23) y cumpliendo (30), dan una correspondencia adecuada entre las expresiones del campo electromagnético clásico y los valores esperados en el K -espacio. Se muestra el caso del momento lineal, para ello se parte de la expresión del momento lineal del campo clásico:

$$\vec{p} = \frac{1}{c^2} \int d\vec{r}^3 \vec{E}(\vec{r}, t) \times \vec{H}(\vec{r}, t) \quad (31)$$

en donde al reemplazar (1) y (2), simplificar y utilizar las expresiones (21) y (22) se obtiene:

$$\vec{p} = \int dk^3 (\hbar \vec{k}) \vec{f}^*(\vec{k}, t) \cdot \vec{f}(\vec{k}, t) \quad (32)$$

En la expresión anterior se identifica al operador multiplicativo momento lineal $\hbar \vec{k}$ en el K -espacio; luego, el valor esperado para cada componente se expresa como:

$$\langle p_\alpha \rangle = \langle S | p_\alpha | S \rangle = \int dk^3 \vec{f}^*(\vec{k}, t) \cdot (\hbar k_\alpha) \vec{f}(\vec{k}, t) \quad (33)$$

2.2 Momento angular del fotón

De manera similar al caso anterior, a partir de la expresión del momento angular del campo electromagnético clásico:

$$\vec{J} = \frac{1}{c^2} \int d\vec{r}^3 \vec{r} \times [\vec{E}(\vec{r}, t) \times \vec{H}(\vec{r}, t)], \quad (34)$$

se obtiene la siguiente expresión:

$$\begin{aligned} \vec{J} = & \sum_{\beta=1}^3 \int dk^3 \vec{f}^*(\vec{k}, t) [-i\hbar \vec{k} \times \nabla_{\vec{k}}] f_\beta(\vec{k}, t) \\ & + \int dk^3 (-i\hbar) \vec{f}^*(\vec{k}, t) \times \vec{f}(\vec{k}, t). \end{aligned} \quad (35)$$

El primer término de la expresión anterior se interpreta como el *valor esperado del operador momento angular orbital* \vec{L} del fotón, debido a su presencia en la forma $-i\hbar \vec{k} \times \nabla_{\vec{k}}$ correspondiente al K -espacio. El

segundo término de (35) es algo nuevo y no es evidente que pueda considerarse como un valor esperado; luego se mostrará que es posible darle la forma de un valor esperado.

3. Estados de polarización del fotón

3.1 Espín del fotón

En esta sección se trabaja con el segundo término de la expresión (35). Primero se consideran dos vectores \vec{A} y \vec{B} cuyas componentes pueden ser números complejos; luego, es posible calcular las componentes del producto $\vec{A}^* \times \vec{B}$ por álgebra vectorial.

La componente k del producto $\vec{A}^* \times \vec{B}$ puede escribirse como:

$$\begin{aligned} (\vec{A}^* \times \vec{B})_k &= (\vec{A}^* \times \vec{B}) \cdot \vec{e}_k = -\vec{A}^* \cdot (\vec{e}_k \times \vec{B}) \\ &= -\vec{A}^* \cdot (E_k \vec{B}) = -\langle \vec{A}, E_k \vec{B} \rangle, \end{aligned} \quad (36)$$

donde E_k es una matriz tal que $E_k \vec{q} = \vec{e}_k \times \vec{q}$ para \vec{q} arbitrario, y los \vec{e}_k son los vectores de la base canónica^[2].

Para calcular los conmutadores de las matrices E_k , notemos que:

$$[E_j, E_k] \vec{q} = E_j E_k \vec{q} - E_k E_j \vec{q} = \vec{e}_j \times (\vec{e}_k \times \vec{q}) - \vec{e}_k \times (\vec{e}_j \times \vec{q}) = (\vec{e}_j \times \vec{e}_k) \times \vec{q},$$

es decir: $[E_1, E_2] \vec{q} = (\vec{e}_1 \times \vec{e}_2) \times \vec{q} = E_3 \vec{q}$ de donde $[E_1, E_2] = E_3$ (pues \vec{q} es arbitrario). Análogamente se obtiene: $[E_2, E_3] = E_1$ y $[E_3, E_1] = E_2$.

Si se define el operador vectorial de tres componentes $\vec{E} \equiv (E_1, E_2, E_3)$, entonces las tres relaciones de conmutación anteriores se pueden escribir como:

$$\vec{E} \times \vec{E} = \vec{E}. \quad (37)$$

Por otro lado a un operador vectorial \vec{V} se lo define como *operador de momento angular*, si cumple con la condición:

$$\vec{V} \times \vec{V} = i\hbar \vec{V}. \quad (38)$$

De (37) se observa que \vec{E} no es un operador vectorial de momento angular; sin embargo, a partir de éste se puede construir un operador \vec{S}

que cumpla (38). Entonces, usando la relación (37) se puede definir el operador $\bar{S} \equiv i\hbar\bar{E}$ que, como puede verificarse directamente, satisface la condición (38).

Las matrices S_k están dadas por:

$$S_1 = i\hbar \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix}, \quad S_2 = i\hbar \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{pmatrix}, \quad S_3 = i\hbar \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \quad (39)$$

Ahora, la j -ésima componente del segundo término de (35) se puede expresar como:

$$\int dk^3 [(-i\hbar) \bar{f}^*(\vec{k}, t) \times \bar{f}(\vec{k}, t)]_j = \langle f | S_j | f \rangle \quad (40)$$

que representa al valor esperado de la componente j del *momento angular intrínseco o espín del fotón*, donde $|f\rangle$ es el vector de estado que representa a la función de estado $\bar{f}(\vec{k}, t)$ y el producto del lado derecho incluye la integración sobre la variable continua \vec{k} .

Se puede verificar que: $\bar{S}^2 = \sum S_j^2 = 2\hbar^2 I$ y además que $[S_j, \bar{S}^2] = 0$, lo cual garantiza la posibilidad de encontrar vectores propios comunes entre el operador \bar{S}^2 y cualquiera de sus componentes S_j . Además, puesto que el operador \bar{S}^2 es proporcional al operador identidad, cualquier vector de estado $|\chi\rangle$ es vector propio de este operador con valor propio $2\hbar^2$; esto indica que el número cuántico de espín del fotón debe ser $s = 1$.

Al resolver la ecuación de valores propios: $S_3|\chi\rangle = \lambda|\chi\rangle$ se obtiene los vectores propios normalizados $|\chi_0\rangle$, $|\chi_+\rangle$ y $|\chi_-\rangle$ asociados a los valores propios 0 , $+\hbar$ y $-\hbar$, respectivamente:

$$|\chi_0\rangle = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}, \quad |\chi_+\rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} e^{-i\pi/4} \\ e^{+i\pi/4} \\ 0 \end{pmatrix}, \quad |\chi_-\rangle = \begin{pmatrix} e^{+i\pi/4} \\ e^{-i\pi/4} \\ 0 \end{pmatrix}. \quad (41)$$

Los vectores propios $|\chi_+\rangle$ y $|\chi_-\rangle$ corresponden a los *estados de polarización circular derecho e izquierdo* cuando \vec{k} es paralelo al eje z .

En el mismo caso (\vec{k} paralelo al eje z), se tiene que:

$$\langle k | \chi_{\pm} \rangle = 0, \quad (42)$$

es decir se cumple la condición de transversabilidad; entonces, estos dos estados son observables (experimentalmente). Para el estado $|\chi_0\rangle$, se tiene que:

$$\langle k | \chi_0 \rangle \neq 0, \quad (43)$$

y es la condición de transversabilidad la que prohíbe (en este caso) la observación experimental de este estado.

4. Conclusiones

La función de onda vectorial en el espacio momentum, $\vec{f}(\vec{k}, t)$, solución de la ecuación (16), queda determinada por el campo eléctrico transformado, ecuación (23). Eligiendo la constante A como en (30), dicha función vectorial justifica su elección como función de estado del fotón, al proporcionar las expresiones correctas para los valores medios clásicos de la densidad de energía electromagnética, ecuación (27), de la densidad de momento lineal, ecuación (33), y de la densidad de momento angular, ecuación (35), correspondientes a los operadores $H \equiv \hbar k c$ (multiplicativo), $P \equiv \hbar \vec{k}$ (multiplicativo) y $L \equiv -i\hbar \vec{k} \times \nabla_{\vec{k}}$ respectivamente.

En la transformada de Fourier del momento angular clásico, ecuación (35), se obtienen dos sumandos; el primero es fácilmente interpretable como el momento angular orbital, pero el segundo resulta ser una agradable sorpresa fruto de haber usado $\vec{f}(\vec{k}, t)$ como función de estado. Dicho sumando da el operador de polarización del fotón, con la característica de transversabilidad de sus vectores propios.

Referencias

- [1] Joseph W. Simmons and Mark J. Guttman, "States, waves and photons: A modern introduction to light", Addison-Wesley Publishing Company, 1970.
- [2] H. G. Valqui, *Apuntes de Mecánica Clásica (Capítulo, Geometría del Cuerpo Rígido)*.

¿Un Experimento crucial para verificar el entanglement de dos fotones?

H.G. Valqui*

RESUMEN

En acuerdo con cierta versión de la Interpretación de Copenhagen, muchos físicos consideran que un fotón carece de un estado definido de polarización, previo a la medición de dicho estado, es decir, antes de haber pasado por un polarizador. Así, cuando fotones gemelos (entangled) son emitidos por una fuente apropiada, cada uno de ellos tiene «todos los posibles estados de polarización» y, si alguno de ellos atraviesa un polarizador, adquiriendo el correspondiente estado de polarización, entonces el otro fotón adquiere instantáneamente el mismo estado de polarización. Este misterioso enlazamiento entre los dos fotones muy separados (por lo cual ya no podrían interactuar entre sí) ha sido supuestamente verificado, entre otros, en los experimentos de A. Aspect.

Pero no todos los físicos (interesados en el tema) se sienten contentos con la interpretación de los resultados de los experimentos de Aspect. Por tal razón – y siguiendo el criterio de «desafío negativo» de K. Popper – aquí presento un experimento crucial para despejar las dudas sobre la existencia del misterioso enlazamiento entre los fotones gemelos.

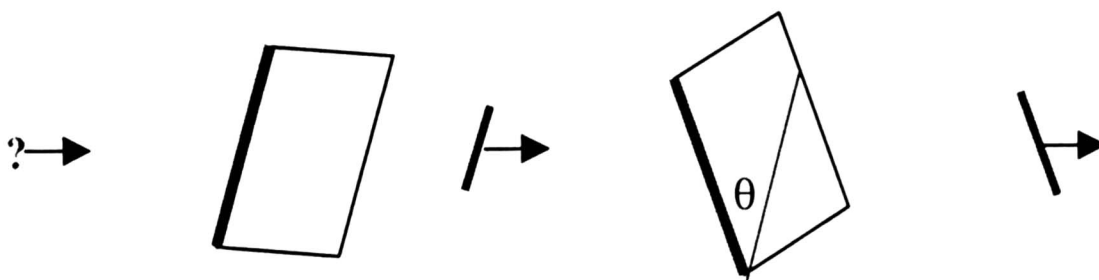
* Facultad de Ciencias - Universidad Nacional de Ingeniería.

ABSTRACT

According to the so called Copenhagen Interpretation many physicists suppose that a photon has no polarization state before such property has been measured, that is, before it has been transmitted by a polarizer. So when twin photons (entangled) are emitted from a suitable source they have «all possible polarization states» and, if one of them is transmitted by a given polarizer, acquiring the corresponding polarization, then his twin photon acquires instantaneously the same polarization state. This entanglement between the twin photons has been supposedly verified, among others, by A.Aspect.

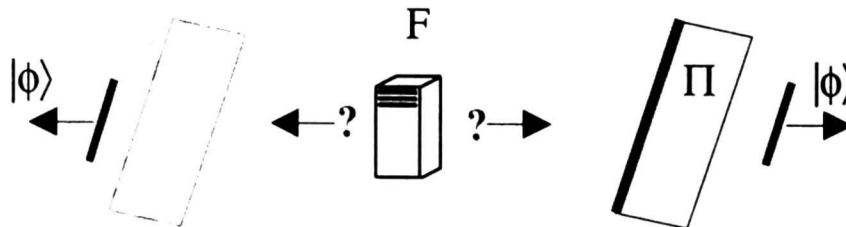
But not every physicist feels happy with the interpretation of Aspect's experimental results. For that reason – and following Popper's falsifiability criterion – I state here a crucial experiment which should clarify any doubt concerning the existence of the entanglement of the separated photons.

1. Es un hecho experimental que si un fotón incide (perpendicularmente) sobre un polarizador, entonces puede atravesarlo o ser absorbido por éste. Además, si lo atraviesa, adquiere un estado de polarización inducido por dicho polarizador. Por otra parte, también se verifica experimentalmente que si un fotón atraviesa consecutivamente dos polarizadores, cuyas direcciones de polarización forman un ángulo θ , entonces la probabilidad de que el fotón pase el segundo polarizador es igual a $\cos^2\theta$ (Ley de Malus)



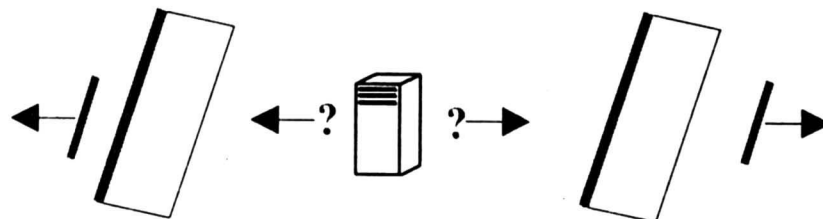
De lo anterior puede verse que si un fotón (aislado) estuviese polarizado (sin que su estado de polarización fuese conocido), sería imposible medir tal estado de polarización. Del hecho que tal fotón atravesase un polarizador solamente podría deducirse que su estado de polarización no era perpendicular a la dirección del polarizador; y si el fotón fuese absorbido, solamente podría deducirse que su polarización no era paralela a la del polarizador.

2. A continuación muestro esquemáticamente un experimento que genera fotones enlazados. Una fuente, F , de átomos de calcio excitados de tal manera que al cesar la excitación sus electrones vuelven al estado fundamental, emitiendo en sentido opuestos fotones gemelos (en cascada), de polarización indefinida, pero conservando un momento angular total nulo.



Cuando el fotón de la derecha atraviesa un polarizador Π , adquirirá un cierto estado de polarización que indicaremos con $|\phi\rangle$. En el mismo instante el fotón de la izquierda –que no atraviesa ningún polarizador– habría adquirido el mismo estado de polarización $|\phi\rangle$.

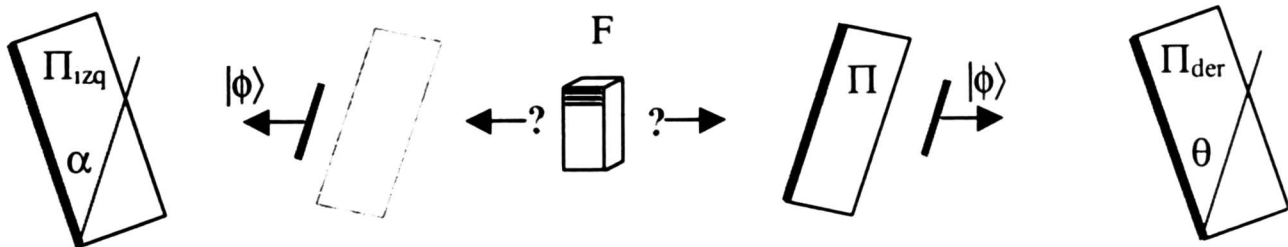
3. ¿Cómo se verifica experimentalmente que los dos fotones adquieren el mismo estado de polarización?



Para ello a los dos fotones se los hace incidir sobre sendos polarizadores «paralelos», obteniéndose que, cada vez, ambos fotones son absorbidos por sus correspondientes polarizadores, o ambos fotones son transmitidos por dichos polarizadores. Según la información de los especialistas, tal coincidencia (ambos absorbidos o ambos transmitidos) es del 100%, para orientaciones arbitrarias de los polarizadores paralelos [Aquí no estoy mencionando algunas consideraciones experimentales importantes, como la garantía de que cada vez se trata de fotones realmente gemelos. Pero tales consideraciones no son significativas para el propósito de este artículo].

4. Para compensar la exigencia formal de un número infinito de experimentos «positivos», necesarios para asegurar la validez de un

«teorema» experimental, el físico-filósofo Karl Popper creó lo que aquí llamo un «desafío negativo» (falsifiability criterion), consistente en la construcción de un experimento que posiblemente muestre resultados que contradigan los resultados ya aceptados.



El experimento propuesto es el siguiente:

- i) Al incidir (solamente) el fotón de la derecha sobre un polarizador Π adquiere un determinado estado de polarización, por ejemplo, $|\phi\rangle$; y el fotón de la izquierda, por efecto del misterioso entrelazamiento, adquiere el mismo estado de polarización $|\phi\rangle$.
- ii) A continuación del polarizador Π colocamos un segundo polarizador Π_{der} , cuya dirección forma un ángulo θ con la dirección de Π , como se muestra en el dibujo. Entonces la probabilidad de que el fotón –en el estado $|\phi\rangle$, porque ya ha sido transmitido por el por el polarizador Π – sea también transmitido por el polarizador Π_{der} será $p_{\text{der}}(\theta) = \cos^2\theta$.
- iii) Después de que el fotón de la izquierda, por efecto del misterioso enlace, haya adquirido el estado $|\phi\rangle$, lo hacemos incidir sobre un polarizador Π_{izq} , cuya dirección forma un ángulo α con el polarizador Π – que este fotón no ha atravesado –, como se muestra en el dibujo. Entonces, la probabilidad de que el fotón (que avanza hacia la izquierda) sea transmitido por el polarizador Π_{izq} será $p_{\text{izq}}(\alpha) = \cos^2\alpha$.
- iv) Las probabilidades $p_{\text{der}}(\theta)$ y $p_{\text{izq}}(\alpha)$ no deberían ser difíciles de medir (para quien ya ha realizado los experimentos que evidencian el misterioso enlace); pero existe una situación que daría resultados más categóricos: elegir los ángulos tales que $\alpha = \theta + \pi/2$, con lo cual debería cumplirse que $p_{\text{der}}(\theta) + p_{\text{izq}}(\alpha) = 1$.
- v) La prueba negativa más contundente se plantea eligiendo $\theta = 0$, con lo cual el fotón de la derecha debería ser siempre transmitido por el polarizador Π_{der} , mientras que el fotón de la

izquierda debería ser siempre absorbido por el polarizador Π_{izq} . Los resultados contrarios sucederían si se eligiese $\theta = \pi/2$; el fotón de la derecha sería siempre absorbido por Π_{der} , mientras que el fotón de la izquierda sería siempre transmitido por Π_{izq} .

vi) ¿Qué sucedería si el fotón de la izquierda no fuese afectado por lo que hubiese sucedido con el fotón de la derecha?

En tal caso, mientras que el fotón de la derecha, después de pasar el polarizador Π ha adquirido el estado $|\phi\rangle$, en cambio, el fotón de la izquierda mantendría un estado de polarización desconocido, lo que posiblemente sería causa de que su transmisión por el polarizador Π_{izq} fuese un número aleatorio, mientras que para el lado derecho se mantendría que $p_{\text{der}}(\theta) = \cos^2(\theta)$.

Conclusión

Puesto que no existen argumentos contra la ley de Malus, el experimento propuesto debería decidir si efectivamente las probabilidades $p_{\text{der}}(\theta)$ y $p_{\text{izq}}(\alpha)$ tienen los valores predichos por los resultados del misterioso entanglement. En caso de no cumplirse la predicción mencionada, habría que aceptar que el fenómeno de entanglement ha sido solamente una persistente ilusión.

Referencias

1. J. Cramer, *Transactional interpretation of quantum mechanics*. *Rev. Mod. Phys* 58,3 July 1986.
2. J. Clauser, A. Shimony, *Bell's theorem: Experimental tests and implications*, *Rep. Prog. Phys.* Vol 41, 1978.
3. B. d'Espagnat, *The Quantum Theory of Reality*, *Scientific Am.* Nov 1979.
4. D. Greenberger, A. Zeilinger, *Quantum theory: still crazy after all these years*, *Physics World*, Sept 1995.
5. A. Aspect, P. Grangier, G. Roger, *Phys. Rev. Lett* 47, 91, 1982.
6. D. Mermin, *Is the Moon there when nobody looks? Reality and quantum theory*, *Physics Today*, April 1985.
7. H. G. Valqui, *El error inicial*, *Rev. ci. uni*, 7, 1, Feb. 2003.

Enlazamiento (Entanglement) clásico

H.G.Valqui*

RESUMEN

La situación de acoplamiento (o entanglement en el lenguaje cuántico) se produce cuando *en un sistema de ecuaciones diferenciales lineales* las funciones en consideración (componentes de una función vectorial o de un producto tensorial) no poseen soluciones independientes una de las otras. Aquí debemos notar ciertas cuestiones:

- i) Las funciones en consideración pueden referirse a situaciones clásicas o a situaciones cuánticas,
- ii) El acoplamiento puede referirse a los estados de ‘movimiento’ de cuerpos físicos, como es el caso de los dos cuerpos; puede referirse a los estados de ciertas propiedades de dos cuerpos físicos, como es el caso de los estados de los espines de dos electrones; o puede referirse a las funciones de estado que son proyecciones de una función de estado, como es el caso del péndulo compuesto; o pueden referirse a las funciones de estado que son factores de una función de estado producto de las funciones anteriores, como es el caso de la membrana vibrante. Posiblemente existan algunos otros casos, la lista anterior no pretende ser exhaustiva.
- iii) La situación de acoplamiento es un fenómeno matemático, que puede corresponder a muy diferentes situaciones físicas. Estas situaciones físicas, a su vez, deberían ser representaciones – más o menos sensatas – de las correspondientes ecuaciones diferenciales lineales. Un caso especial, por razones netamente físicas, surge en el modelo cuántico de las partículas idénticas, sin paralelo en el modelo clásico.

* Facultad de Ciencias - Universidad Nacional de Ingeniería

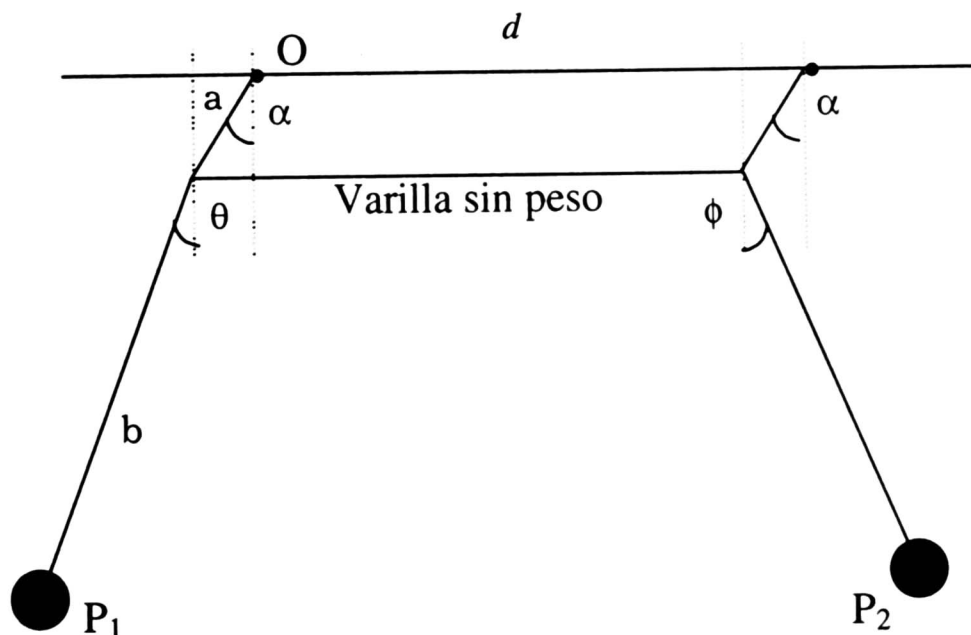
- iv) Las ecuaciones diferenciales lineales en cuestión pueden corresponder a un modelo clásico o a un modelo cuántico. Matemáticamente no hay diferencias entre los dos casos; físicamente las diferencias pueden ser notables. Por otra parte es necesario tener presente las diferencias entre los casos de funciones ordinarias de una variable y funciones de varias variables.
- v) Estas aclaraciones son necesarias, entre otras cosas, para evitar interpretaciones físicas ‘demasiado profundas’ que se constituyan en una camisa de fuerza para las verificaciones experimentales del caso.

ABSTRACT

The question of entanglement appears when functions which are solutions of *a system of linear differential equations*, are not independent of each other. In that sense we should note the following:

- i) The above mentioned functions may refer to a physical problem described classically, or described by the model of the quantum mechanics.
- ii) The entanglement may refer to the ‘motion’ of some physical particles, as in the problem of two bodies; or it may refer to the states of some properties of two physical particles, as in the case of the spins of two electrons; or it may refer to the function states which are projections of another state function, as is the case with the compound pendulum; or it may refer to state functions which are factors of a state function which is product of such function, as in the case of vibrating membrane. This list is not exhaustive; possibly other cases of entanglement may exist.
- iii) Entanglement or coupling is a mathematical fact, which can correspond to very different physical phenomena. On the other side, such physical phenomena should be – more or less – sensible representations of the corresponding linear differential equations. A special case – based on genuine physical motives – appears in the quantum model for identical particles, which has no analogue in the classical model.
- iv) The considered linear differential equations may arise from a classical model or a quantum one. Mathematically there is no difference between both models; physically there are big ones. But there are differences between the situations described recurring to ordinary functions of one variable and function of several variables.
- v) The above exposed considerations are important, besides other questions, in order to avoid ‘profound physical interpretations’ which may become a strait jacket for the interpretations of the experimental facts.

I) Péndulos acoplados



$$\mathbf{r}_1 = (a \sin \alpha + b \sin \theta, a \cos \alpha + b \cos \theta) \quad \mathbf{r}_2 = (d + a \sin \alpha + b \sin \phi, a \cos \alpha + b \cos \phi)$$

En la figura se muestran dos péndulos acoplados por una varilla ideal. Aplicando la Segunda Ley de Newton para el movimiento de las peonzas, obtenemos las ecuaciones para las tres variables angulares, donde los «puntitos» superiores indican derivación temporal:

$$g \sin \theta = - (a \ddot{\alpha} + b \ddot{\theta}) \cos \theta + (a \cos \alpha + b \cos \theta) \ddot{\alpha} \sin \theta \quad [01]$$

$$g \sin \phi = - (a \ddot{\alpha} + b \ddot{\phi}) \cos \phi + (a \cos \alpha + b \cos \phi) \ddot{\alpha} \sin \phi \quad [02]$$

$$2g \sin \alpha = - (2a \ddot{\alpha} + b \ddot{\theta} + b \ddot{\phi}) \cos \alpha + (2a \cos \alpha + b \cos \theta + b \cos \phi) \ddot{\alpha} \sin \alpha \quad [03]$$

Estas ecuaciones del movimiento son ecuaciones diferenciales **no lineales**, cuyas soluciones deben ser calculadas numéricamente. Semejantemente al caso del péndulo simple, realizaremos algunas simplificaciones que las conviertan en **ecuaciones lineales**. Supondremos (y las soluciones valdrán en tales condiciones) que los ángulos de desviación son bastante pequeños como para considerar insignificantes los efectos de, por ejemplo, θ^2 , $\theta \dot{\theta}$, α^2 , $\theta \dot{\alpha}$, etc. Entonces, las ecuaciones [01], [02] y [03] toman las formas (lineales):

$$g \sin \theta + (a \ddot{\alpha} + b \ddot{\theta}) \cos \theta = 0 \quad [04]$$

$$g \phi + (a \alpha + b \phi)'' = 0 \quad [05]$$

$$2g \alpha + (2a \alpha + b \theta + b \phi)'' = 0 \quad [06]$$

$$\text{de donde, } [04] + [05] - [06], \text{ resulta que } \theta + \phi = 2 \alpha \quad [07]$$

Eliminando la variable α , las ecuaciones de estado serán:

$$2g \theta + (a + 2b)\theta'' + a \phi'' = 0 \quad [08]$$

$$2g \phi + a \theta'' + (a + 2b) \phi'' = 0 \quad [09]$$

o, en forma matricial:

$$2g \begin{bmatrix} \theta \\ \phi \end{bmatrix} + \begin{bmatrix} a+2b & a \\ a & a+2b \end{bmatrix} \begin{bmatrix} \theta'' \\ \phi'' \end{bmatrix} = \mathbf{0}, \text{ es decir } \begin{bmatrix} \theta \\ \phi \end{bmatrix}'' + g/[2b(a+b)] \begin{bmatrix} a+2b & -a \\ -a & a+2b \end{bmatrix} \begin{bmatrix} \theta \\ \phi \end{bmatrix} = \mathbf{0}$$

Designando con $\chi \equiv [\theta, \phi]$ al vector (columna) de estado del sistema (que describe el movimiento de los péndulos), la ecuación anterior puede escribirse

$$\chi'' + g/[2b(a+b)] \Omega \chi = \mathbf{0} \quad [10]$$

El hecho de que la matriz Ω no sea diagonal significa que los movimientos de los péndulos están acoplados o enlazados. Veamos qué condiciones se requieren para desacoplar o desenlazar dichos movimientos.

Los vectores propios de Ω son $\mathbf{u}_1 = [1 \ -1]$ y $\mathbf{u}_2 = [1 \ 1]$, con los correspondientes valores propios $\lambda_1 = 2(a+b)$, $\lambda_2 = 2b$.

Como los vectores propios constituyen una base (ortogonal) para describir cualquier estado (o solución de la ecuación de estados), podemos escribir

$$\chi = q_1 \mathbf{u}_1 + q_2 \mathbf{u}_2, \quad [11]$$

lo que reemplazado en [10], da

$$q_1'' \mathbf{u}_1 + q_2'' \mathbf{u}_2 + g/[2b(a+b)] (q_1 \lambda_1 \mathbf{u}_1 + q_2 \lambda_2 \mathbf{u}_2) = \mathbf{0}$$

lo que permite escribir las dos ecuaciones desacopladas:

$$q_1'' + w_1^2 q_1 = 0, \quad q_2'' + w_2^2 q_2 = 0 \quad [12]$$

$$\text{donde } w_1^2 = g/b \quad w_2^2 = g/(a+b) = w_1^2/(1 + a/b) \quad [13]$$

Entonces, $q_k(t) = A_k \sin(w_k t + s_k)$, $k = 1, 2$; donde A_k , s_k son constantes de integración, determinadas por las condiciones iniciales. Ahora, de [11] obtenemos que,

$$[\theta, \phi] = [q_1 + q_2, -q_1 + q_2] \quad [14]$$

es decir,

$$\theta(t) = A_1 \sin(w_1 t + s_1) + A_2 \sin(w_2 t + s_2) \quad [15A]$$

$$\phi(t) = -A_1 \sin(w_1 t + s_1) + A_2 \sin(w_2 t + s_2) \quad [15B]$$

donde se hace explícita la forma de enlace de las soluciones del movimiento: El movimiento de cada uno de los péndulos depende de las dos frecuencias w_1 , w_2 .

En esta aproximación lineal, la energía del sistema tiene la forma:

$$(m/2) (a \dot{\alpha} + b \dot{\theta})^2 + (m/2) (a \dot{\alpha} + b \dot{\phi})^2 - 2mg(a + b) = E \quad [16]$$

es decir,

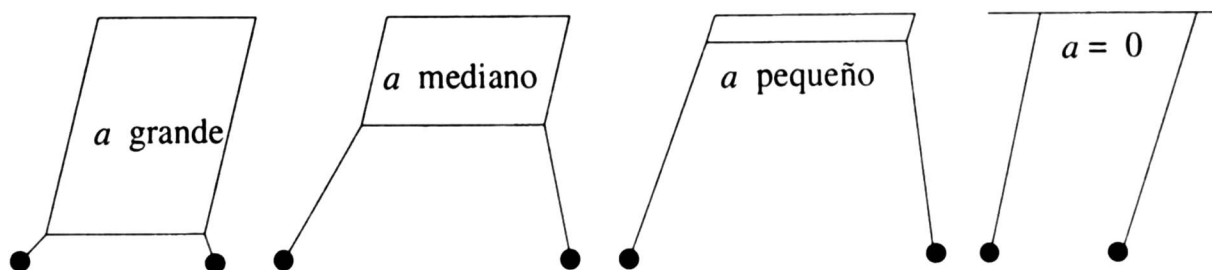
$$[(a/2 + b)^2 + (a/2)^2] \dot{\theta}^2 + [(a/2)^2 + (a/2 + b)^2] \dot{\phi}^2 + a(a + 2b) (\dot{\theta} \dot{\phi}) - 4 \cdot g(a + b) = 2E/m \quad [17]$$

Esta ecuación [17], junto con las ecuaciones [15], muestra que los dos péndulos intercambian energía cinética (cuando $\dot{\theta}^2$ crece, entonces $\dot{\phi}^2$ decrece, y recíprocamente).

La frecuencia del intercambio de energía se obtiene de $w_1 - w_2 = w_1 [1 - 1/(1 + a/b)^{1/2}]$, es decir,

$$w_1 - w_2 = (a/b) w_1 / [1 + a/b + (1 + a/b)^{1/2}] \quad [18]$$

donde puede apreciarse cuando $a \rightarrow 0$ el tiempo de transferencia se vuelve infinito, lo que corresponde al desacoplamiento de los dos péndulos (uno actúa independiente del otro).



El intercambio de energía entre los dos péndulos se realiza a través de la varilla horizontal, que se encuentra a la distancia a del soporte. Para valores grandes de a el intercambio es rápido. Para a pequeño el intercambio es lento, tendiendo a infinito cuando el valor de a tiende a cero. Esto puede ser claramente observado experimentalmente (donde debe tenerse cuidado de que la varilla sea bastante liviana; en tal sentido conviene usar, por ejemplo, un sorbete de líquidos).

Con el objeto de apreciar matemáticamente el desacoplamiento, escribiremos las expresiones [15A, B] vectorialmente:

$$[\theta(t), \phi(t)] = A_1 \sin(w_1 t + s_1) \begin{bmatrix} 1 & -1 \end{bmatrix} + A_2 \sin(w_2 t + s_2) \begin{bmatrix} 1 & 1 \end{bmatrix} \quad [15C]$$

Para el caso de $a = 0$, es decir, $w_2 = w_1 \equiv w$, las funciones de estado, $\theta(t)$, $\phi(t)$, resultan desacopladas,

$$\begin{bmatrix} \theta(t) \\ \phi(t) \end{bmatrix} = A_3 \sin(w t + s_3) \begin{bmatrix} 1 \\ 0 \end{bmatrix} + A_4 \sin(w t + s_4) \begin{bmatrix} 0 \\ 1 \end{bmatrix} \quad [19]$$

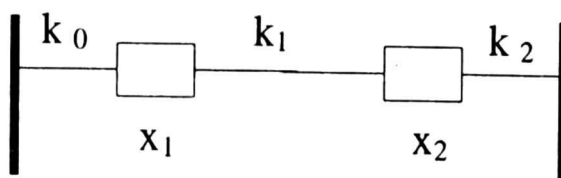
$$\begin{aligned} \text{donde, } A_3 \sin(wt + s_3) &= A_1 \sin(wt + s_1) + A_2 \sin(wt + s_2) \\ A_4 \sin(wt + s_4) &= -A_1 \sin(wt + s_1) + A_2 \sin(wt + s_2) \end{aligned}$$

Lo que muestra que las funciones de estados, $\theta(t)$, $\phi(t)$, son independientes entre sí.

II) Resortes acoplados

Consideremos dos bloques de masas m , conectados a 3 resortes de constantes elásticas de valor k_i , como se muestra, donde los extremos se encuentra a la distancia L .

$$z_1 = x_1 - x_{10}, \quad z_2 = x_2 - x_{20}$$



Las longitudes de los resortes son:
 x_{10} , $x_{20} - x_{10}$, $L - x_{20}$

Aplicando la segunda ley de Newton se obtienen las ecuaciones de los movimientos:

$$-k_0 z_1 + k_1 (z_2 - z_1) = m z_1''$$

$$-k_1 (z_2 - z_1) - k_2 z_2 = m z_2'' \quad [20]$$

o, en forma matricial,

$$\begin{bmatrix} -A & C \\ C & -B \end{bmatrix} \begin{bmatrix} z_1 \\ z_2 \end{bmatrix} = \begin{bmatrix} z_1'' \\ z_2'' \end{bmatrix} \quad \begin{aligned} A &= (k_0 + k_1)/m \\ B &= (k_2 + k_1)/m \\ C &= k_1/m \end{aligned}$$

o también,

$$\Omega z = z'' \quad [21]$$

Los vectores propios (no unitarios) de Ω son $\mathbf{u}_k = [C \quad \lambda_k + A]$ o, si se prefiere, $[\lambda_k + B \quad C]$, para $k = 1, 2$; con $2\lambda_1 = -A - B + R$, $2\lambda_2 = -A - B - R$, donde $R = [(A - B)^2 + 4C^2]^{1/2}$

Puede verificarse que estos valores propios son negativos.

Ahora los vectores propios (no unitarios) ortogonales se pueden escribir como

$$\mathbf{u}_1 = [B - A + R \quad C] \quad \mathbf{u}_2 = [C \quad -(B - A + R)] \quad [22]$$

que usamos como base para expresar el vector solución

$$\mathbf{z} = q_1 \mathbf{u}_1 + q_2 \mathbf{u}_2$$

lo cual reemplazado en [21] da, como anteriormente, las ecuaciones $q_1'' + [(A + B - R)/2] q_1 = 0$, $q_2'' + [A + B + R]/2 q_2 = 0$ [23] cuyas soluciones (salvo un factor constante) son

$$q_1(t) = \text{sen}(w_1 t + \alpha_1) \quad q_2(t) = \text{sen}(w_2 t + \alpha_2) \quad [24]$$

conocidas como coordenadas normales. Se trata de las coordenadas del sistema físico constituido por los dos bloques, y no de las coordenadas de los bloques. Aquí

$$w_1 = [(A + B - R)/2]^{1/2}, \quad w_2 = [(A + B + R)/2]^{1/2}$$

A los vectores-solución

$$\mathbf{Z}_1(t) = \text{sen}(w_1 t + \alpha_1) \mathbf{u}_1 \quad \mathbf{Z}_2(t) = \text{sen}(w_2 t + \alpha_2) \mathbf{u}_2 \quad [25]$$

los designamos como (soluciones o) estados propios del sistema físico. En cada estado propio, los bloques se mueven con la misma frecuencia y fase, pero con diferentes amplitudes (determinadas por las componentes del correspondiente vector propio). El estado general del movimiento será una combinación lineal de los estados propios:

$$\mathbf{z}(t) = C_1 \mathbf{Z}_1(t) + C_2 \mathbf{Z}_2(t) \quad [26]$$

En [25] puede apreciarse que los estados propios del sistema son independientes (por ello dichos estados son importantes en la descripción del movimiento del sistema); pero lo que físicamente nos interesa es el acoplamiento entre los movimientos de los bloques. En [26] puede verse que los movimientos de los bloques, dados por $z_1(t)$ y $z_2(t)$, se encuentran acoplados (o enlazados).

En el dibujo puede apreciarse que el enlace está determinado por la constante k_1 , es decir, por la constante C .

Para $k_1 = 0$ ($C = 0$) los bloques independizan sus movimientos. En efecto, para

$$k_1 = 0, \quad A > B \text{ se obtiene,}$$

$$w_1 = B^{1/2}, \quad w_2 = A^{1/2}, \quad \mathbf{u}_1 = 2B [1 \quad 0], \quad \mathbf{u}_2 = -2B [0 \quad 1]$$

con lo cual, en esta situación particular, $\mathbf{Z}_1(t)$ describe el movimiento del primer bloque, mientras que $\mathbf{Z}_2(t)$ describe el movimiento del segundo bloque; siendo cada uno de los movimientos independiente del otro.

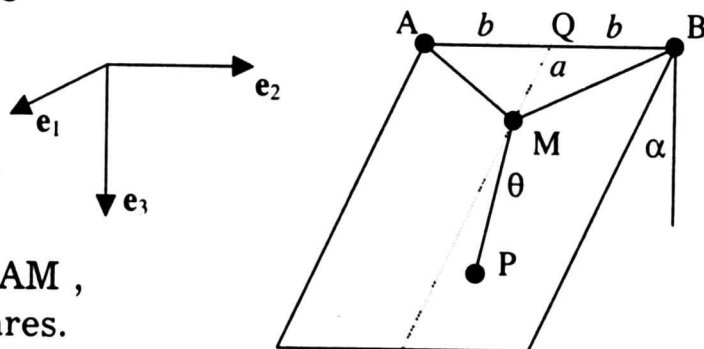
Por otra parte, de [20] se obtiene la ecuación de la conservación de la energía:

$(m/2) \dot{z}_1^2 + (k_0/2) z_1^2 + (m/2) \dot{z}_2^2 + k_2 z_2^2 + (k_1/2)(z_2 - z_1)^2 = E$

donde puede apreciarse que el quinto sumando de la izquierda es la energía de acoplamiento; la misma que tiende a cero cuando tiende a anularse el enlace entre los dos bloques, por tender k_1 a cero.

III) El péndulo compuesto

El péndulo compuesto es un sistema físico formado por un único péndulo ABMP, donde se verifica que las cuerdas ideales AM, MB y MQ permanecen coplanares. El ángulo α indica cuánto se aparta el plano ABM de la vertical, mientras que el ángulo θ corresponde al 'péndulo' MP. Aplicando la Segunda Ley a la peonza en P, obtenemos las ecuaciones **no lineales**:



$\alpha > 0$ en sentido de e_1

$\theta > 0$ en sentido de e_2

$$a \alpha'' + \ell \alpha'' \cos \theta - 2 \ell \theta' \alpha' \sin \theta = -g \sin \alpha \quad [27]$$

$$\ell \theta'' + (a + \ell \cos \theta) \alpha'^2 \sin \theta + g \cos \alpha \sin \theta = 0 \quad [28]$$

Como en el caso de los péndulos acoplados, consideraremos valores de los ángulos, θ , α , suficientemente pequeños, de manera que θ^2 , α^2 , $\theta \alpha$, θ'^2 , α'^2 , y productos del mismo orden, puedan ser considerados insignificantes. Entonces las ecuaciones [27] y [28] se reducen a dos ecuaciones lineales,

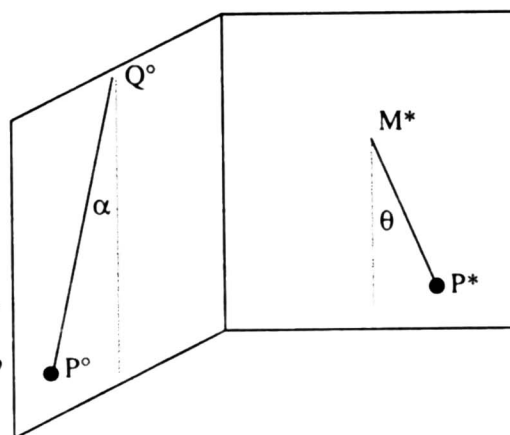
$$(a + \ell) \alpha'' + g \alpha = 0 \quad [29]$$

$$\ell \theta'' + g \theta = 0 \quad [30]$$

que son las ecuaciones de dos péndulos simples, independientes entre sí, de longitudes $a + \ell$ y ℓ , respectivamente.

En este ejemplo, podemos apreciar el interesante caso de que, a pesar de tratarse de un único péndulo (no simple), el sistema físico se comporta como si se tratase de dos péndulos simples (desacoplados) independientes uno del otro.

El comportamiento de los 'dos péndulos' desacoplados se puede verificar experimentalmente, montando ortogonalmente dos péndulos simples, uno de longitud $\alpha + \ell$, en un plano paralelo al plano XOZ, cuyo soporte Q° está en la prolongación de AB; el otro, de longitud ℓ , en un plano paralelo al YOZ, cuyo soporte M^* se encuentra a la altura de M (este punto, en la aproximación usada, mantiene una altura constante).



A continuación, dando los valores iniciales adecuados, se puede lograr que los dos péndulos ortogonales se muevan isócronamente con cada una de las correspondiente proyecciones del péndulo compuesto. También se puede lograr el efecto contrario (lo cual es una tarea que debe ser realizada cuidadosamente y con mucha paciencia): Estando el péndulo compuesto detenido, y los péndulos proyecciones en movimiento, se puede dar al péndulo compuesto las condiciones iniciales tales que permitan que las proyecciones de su movimiento sean isócronas con los movimientos de los péndulos simples ortogonales.

IV) El problema de dos cuerpos

Sean dos cuerpos, uno de masa m_1 en P_1 , y otro de masa m_2 en P_2 , que interactúan según las fuerzas

$$\mathbf{F}_{12} = -K (\mathbf{r}_1 - \mathbf{r}_2) / |\mathbf{r}_1 - \mathbf{r}_2|^3 \quad [31]$$

sobre el primero; \mathbf{F}_{21} sobre el segundo, donde $\mathbf{F}_{12} + \mathbf{F}_{21} = 0$.

Entonces, en una referencia inercial se cumplirá

$$\mathbf{F}_{12} = m_1 d^2 \mathbf{r}_1 / dt^2 \quad \mathbf{F}_{21} = m_2 d^2 \mathbf{r}_2 / dt^2 \quad [32]$$

donde resulta evidente que los movimientos de los dos cuerpos se encuentran acoplados o enlazados.

Si ahora consideramos el vector de posición del centro de masa, $\mathbf{R} = (m_1 \mathbf{r}_1 + m_2 \mathbf{r}_2) / (m_1 + m_2)$, y definimos el *vector de posición relativa* del cuerpo 1 con respecto al cuerpo 2, como $\mathbf{r} = \mathbf{r}_1 - \mathbf{r}_2$, obtenemos que

$$\mathbf{r}_1 = \mathbf{R} + m_2 / M \mathbf{r}, \quad \mathbf{r}_2 = \mathbf{R} - m_1 / M \mathbf{r}$$

Sumando y restando las ecuaciones [32] obtenemos las ecuaciones diferenciales para \mathbf{R} y para \mathbf{r} ,

$$d^2\mathbf{R}/dt^2 = 0 \quad d^2\mathbf{r}/dt^2 = m d^2\mathbf{r}/dt^2, \text{ donde } m = m_1m_2/M \quad [33]$$

que son dos ecuaciones completamente independientes entre sí: Es decir, el movimiento del centro de masa, $\mathbf{R}(t)$, se realiza independientemente del movimiento relativo, $\mathbf{r}(t)$: Los estados (de movimiento) de los sistemas descritos están totalmente desacoplados. Debemos destacar que no se trata de los estados de las dos partículas (que sí se encuentran acoplados), sino de los estados de dos sistemas físico-matemáticos. Esto es parcialmente similar al caso del péndulo compuesto, donde las ecuaciones [29] y [30] se refieren a dos péndulos que existen matemáticamente, pero no físicamente.

Finalmente debemos tener presente que si, en el movimiento relativo (que se desarrolla en un plano), queremos separar el movimiento radial $r(t)$, del movimiento angular $q(t)$, dichos movimientos resultan acoplados por la conservación del momento angular,

$$r^2 d\theta/dt = \text{constante} \quad [34]$$

V) La membrana vibrante

La función de estado, $\psi(x, z, t)$, que describe las vibraciones de una membrana debe satisfacer la ecuación de onda,

$$c^2 \nabla^2 \psi - \partial^2 \psi / \partial t^2 = 0 \quad [35]$$

donde debe suponerse que las vibraciones resultantes son libres, generadas solamente por las condiciones iniciales, pero sometidas a ciertas condiciones de contorno (como sucede en el caso de la ecuación de Schroedinger). Aquí no nos interesa hallar la solución de la ecuación de onda en forma compacta, sino en su forma de producto de funciones.

Ensayando la separación de variables $\psi(x, z, t) = U(x, z) T(t)$ obtenemos $(\nabla^2 U)/U - (d^2T/d^2t)/c^2T = 0$. Aquí ambos sumandos son independientes, por lo cual cada uno de ellos debe ser igual a una misma constante. Con el objeto de que las soluciones se mantengan acotadas elegimos que dicha constante sea negativa, $-k^2$; así obtenemos las ecuaciones:

$$\nabla^2 U + k^2 U = 0, \quad d^2T/d^2t = w^2T, \text{ con } w = k/c \quad [36]$$

siendo k una constante arbitraria. Es decir, en realidad se trata de dos familias mono-paramétricas de ecuaciones diferenciales, donde el producto de cada par de soluciones es, a su vez, una solución particular de la ecuación [35].

La solución de la segunda ecuación [36] es directa,

$$T_k(t) = A_k e^{i \omega t} + B_k e^{-i \omega t} \quad [37]$$

Para hallar las soluciones de la segunda ecuación, consideraremos el caso en el que el contorno es circular y permanece fijo.

$$U(x, z) = 0 \quad \text{para } x^2 + z^2 = r^2 \quad [38]$$

Habiendo tomado el centro de la membrana no deformada como el origen de coordenadas, conviene, por razones de simetría, trabajar con coordenadas polares, r, θ ; entonces la primera ecuación diferencial de [36] toma la forma.

$$(1/r) [\partial/\partial r (r \partial U/\partial r)] + (1/r) \partial^2 U/\partial^2 \theta + k^2 U = 0$$

o también,

$$\partial^2 U/\partial r^2 + (1/r) \partial U/\partial r + (1/r) \partial^2 U/\partial^2 \theta + k^2 U = 0 \quad [39]$$

Ahora ensayamos la separación de variables $U(r, \theta) = R(r) Y(\theta)$, con lo cual se obtienen dos ecuaciones diferenciales,

$$R'' + (1/r) R' + (k^2 - v^2/r^2) R = 0, \quad Y'' + v^2 Y = 0 \quad [40]$$

donde la constante de separación, v^2 , ha sido elegida de manera que las soluciones se mantengan acotadas.

Aquí queremos destacar varias cosas:

- i) Las soluciones R, Y de las ecuaciones [40] no sólo están acopladas por el parámetro v , sino también por el parámetro k ; de manera que la escritura correcta de dichas ecuaciones debería ser:

$$R_{kv}'' + (1/r) R_{kv}' + (k^2 - v^2/r^2) R_{kv} = 0, \quad Y_{kv}'' + v^2 Y_{kv} = 0 \quad [41]$$

- ii) La solución y del problema planteado por la ecuación [35] **y las condiciones de contorno que se deseen estipular**, no puede ser (excepto en algunos casos particulares muy especiales) igual al producto $R_{kv} Y_{kv} T_k$, sino las combinaciones lineales adecuadas,

$$\psi(r, \theta, t) = \sum_{kv} M_{kv} R_{kv}(r) Y_{kv}(\theta) T_k(t) \quad [42]$$

donde las constantes M_{kv} deben ser determinadas por las condiciones de contorno. Aquí podemos apreciar claramente que las funciones factores, R, Y, T se encuentran inevitablemente acopladas.

Suponiendo (como es usual) que se desee que el borde, para $r = a$, se mantenga fijo, entonces deberá cumplirse que, $\psi(a, \theta, t) = 0$, es decir, las constantes M_{kv} deben satisfacer la ecuación

$$0 = \sum_{kv} M_{kv} R_{kv}(a) Y_{kv}(\theta) T_k(t), \quad \text{para todo } \theta, t \quad [43]$$

- iii) En el caso muy particular en que las funciones factores no se encontrasen acopladas, es decir, la solución sea de la forma,

$$\psi(r, \theta, t) = R_{kv}(r) Y_{kv}(\theta) T_k(t) \quad [44]$$

para ciertos valores fijos de los números k, v , entonces las condiciones iniciales deben ser estipuladas de manera que se cumpla

$$\psi(r, \theta, 0) = R_{kv}(r) Y_{kv}(\theta) T_k(0) \text{ para todo valor de } r \text{ y de } \theta \quad [45a]$$

$$0 = R_{kv}(a) Y_{kv}(\theta) T_k(t), \text{ para todo } \theta, t \quad [45b]$$

cuya única solución es la solución trivial, $\psi = 0$.

Conclusiones

- i) Tanto en el caso de los dos resortes acoplados, como en el caso de la aproximación lineal de los dos péndulos acoplados se ha mostrado el parámetro de acoplamiento, el mismo que al tender a cero produce el desacoplamiento.
- ii) En el caso del péndulo compuesto se ha mostrado cómo un sistema físico, con componentes físicas no desacoplables, puede tener (en su aproximación lineal) dos componentes o proyecciones totalmente desacopladas.
- iii) En el caso del Problema de los Dos Cuerpos se ha mostrado que dos cuerpos acoplados por un potencial de interacción, pueden ser tratados como dos sistemas matemáticamente desacoplados (el de Centro de masa y el de Movimiento relativo).
- iv) En el caso de la cuerda vibrante se ha mostrado cómo una función de estado puede ser expresada como el producto de funciones factores, las mismas que, en este caso, resultan inevitablemente acopladas.
- v) Posiblemente existan otros casos de acoplamientos clásicos, que pongan en evidencia características interesantes de tal situación.

Dominios de Convergencia de Series de Potencias de Varias Variables

*Renato Benazic Tomé**

RESUMEN

En el presente trabajo, estudiamos el conjunto de puntos de \mathbb{C}^n para los cuales una serie de potencia de varias variables complejas dada es convergente. Probaremos que este conjunto es un Dominio de Reinhardt completo con base logarítmicamente convexa.

* *Universidad Nacional de Ingeniería Instituto de Matematica y Ciencias Afines - IMCA
Universidad Nacional Mayor de San Marcos Facultad de Ciencias Matematicas*

1. Introducción

En la Teoría de Funciones de una variable compleja, es bien conocido el siguiente resultado [2]: Dada la serie de potencias de una variable compleja centrada en $z_0 \in \mathbb{C}$

$$\sum_{n=0}^{\infty} a_n (z - z_0)^n,$$

existe un único número real $R \in [0, +\infty]$ tal que:

- (1) La serie de potencias converge absolutamente en $D_R(z_0) \subseteq \mathbb{C}$.
- (2) La serie de potencias diverge en $\mathbb{C} - D_R[z_0]$
- (3) Si $0 < r < R$ entonces la serie converge uniformemente en

$$D_r(z_0) \subseteq \mathbb{C}$$

en donde $D_R(z_0)$ (respectivamente $D_R[z_0]$) denota al disco abierto (respectivamente, disco cerrado) de \mathbb{C} centrado en z_0 y de radio R . El número R es llamado *radio de convergencia* de la serie de potencias dada. Además, por el Teorema de Cauchy - Hadamard, este radio de convergencia puede ser calculado por la fórmula

$$\frac{1}{R} = \limsup_{n \rightarrow \infty} |C_n|^{\frac{1}{n}}.$$

El disco abierto $B_r(0)$ es llamado *dominio de convergencia* de la serie de potencias.

Recordemos además que en el círculo $S_R[z_0] = \partial B_R(z_0)$ la convergencia de la serie de potencias es, por lo general muy complicada.

Una pregunta natural sería ¿Qué ocurre si en lugar de trabajar con series de una variable, consideramos series de potencias de varias variables? Las series de potencias de varias variables ocupan un papel fundamental en la teoría de funciones analíticas de varias variables complejas (el lector interesado puede revisar las referencias [4], [3], [5]). Por esta razón es importante conocer como son los subconjuntos de \mathbb{C}^n para los cuales una serie de potencia de varias variables sea convergente.

En el presente trabajo respondemos a la interrogante planteada, caracterizando a los dominios de convergencia de una serie de potencias de varias variables complejas.

2. Algunos conceptos preliminares

En lo que sigue, \mathbb{C}^n denotará al conjunto de todas las n -uplas $z = (z_1, z_2, \dots, z_n)$ de números complejos. Un *polidisco abierto* (respectivamente, *polidisco cerrado*) en \mathbb{C}^n , de centro $a = (a_1, a_2, \dots, a_n) \in \mathbb{C}^n$ y *poliradio* $r = (r_1, r_2, \dots, r_n) \in (\mathbb{R}^+)$ (\mathbb{R}^+ denota al conjunto de los números reales positivos), denotado por $\Delta(a, r)$ (respectivamente, $\Delta[a, r]$; es el conjunto definido por

$$\begin{aligned}\Delta(a, r) &= \{z = (z_1, z_2, \dots, z_n) \in \mathbb{C}^n: |z_j - a_j| < r_j, \forall 1 \leq j \leq n\} \\ \Delta[a, r] &= \{z = (z_1, z_2, \dots, z_n) \in \mathbb{C}^n: |z_j - a_j| \leq r_j, \forall 1 \leq j \leq n\}\end{aligned}$$

Observe que

$$\Delta(a, r) = Dr_1(a_1) \times Dr_2(a_2) \times \dots \times Dr_n(a_n) \text{ y } \Delta[a, r] = Dr_1[a_1] \times Dr_2[a_2] \times \dots \times Dr_n[a_n]$$

En algunos casos conveniente considerar *polidiscos extendidos*, es decir polidiscos en que alguna (o toda) componente de su poliradio sea ∞ .

A lo largo del presente trabajo, usamos la notación de los multi-índices de Schwarz: Un *multi-índice n -dimensional* Q es una n -upla de números enteros no negativos. La *norma* $|Q|$ del multi-índice $Q = (q_1, q_2, \dots, q_n)$ se define como $|Q| = q_1 + q_2 + \dots + q_n$. El *factorial* $Q!$ del multi-índice $Q = (q_1, q_2, \dots, q_n)$ se define por $Q! = q_1! q_2! \dots q_n!$. Sean $Q = (q_1, q_2, \dots, q_n)$ y $\alpha = (\alpha_1, \alpha_2, \dots, \alpha_n)$ dos multi-índices, escribiremos $\alpha \leq Q$ si y sólo si $\alpha_j \leq q_j$, para todo $1 \leq j \leq n$. Significado análogo tiene las notaciones $\alpha < Q$, $\alpha \geq Q$ y $\alpha > Q$. Además, si $z = (z_1, z_2, \dots, z_n) \in \mathbb{C}^n$ y $Q = (q_1, q_2, \dots, q_n)$ es un multi-índice, denotamos

$$z^Q = z_1^{q_1} z_2^{q_2} \dots z_n^{q_n}.$$

Una *multi-sucesión* de números reales o complejos es una función cuyo dominio es el conjunto de multi-índices y su contradominio es el conjunto de los números reales o complejos, en símbolos

$$\begin{aligned}a: \quad \mathbb{N}^n &\rightarrow K \\ Q &\mapsto a(Q) = a_Q = a_{q_1, \dots, q_n}\end{aligned}$$

en donde \mathbb{N} denota al conjunto de los enteros no negativos y $K = \mathbb{R}$ ó \mathbb{C} .

Decimos que $c \in K$ es el *límite* de la multi-sucesión (a_Q) si y sólo si dado un $\varepsilon > 0$, existe un $N \in \mathbb{N}$ tal que $|a_{q_1, \dots, q_n} - c| < \varepsilon$ siempre que

$q_1, \dots, q_n > N$. Una multi-sucesión es llamada *convergente* cuando tiene límite, en caso contrario la multi-sucesión es llamada *divergente*.

El lector puede verificar sin dificultad alguna que dada dos multi-sucesiones convergentes entonces su suma, resta, multiplicación y división (siempre que el cociente sea distinto de cero) son convergentes.

A toda multi-sucesión (a_Q) en K , se le asocia la *multi-sucesión de sumas parciales*

$$s: \mathbb{N}^n \rightarrow K$$

$$Q \mapsto s(Q) = \sum_{\alpha \leq Q} a_\alpha = \sum_{\alpha_1=1}^{q_1} \cdots \sum_{\alpha_n=1}^{q_n} a_{\alpha_1, \dots, \alpha_n}$$

Decimos que la *multi-serie* $\sum_Q a_Q$ es *convergente* si y sólo si la multi-sucesión de sumas parciales (s_Q) es convergente, es decir, existe un $c \in K$ con la propiedad de que dado un $\varepsilon > 0$, existe un $N \in \mathbb{N}$ tal que

$$\left| \sum_{\alpha_1=1}^{q_1} \cdots \sum_{\alpha_n=1}^{q_n} a_{\alpha_1, \dots, \alpha_n} - c \right| < \varepsilon$$

siempre que $q_1, \dots, q_n > N$.

Decimos que la multi-serie $\sum_Q a_Q$ es *absolutamente convergente* si y sólo si $\sum_Q |a_Q|$ es convergente.

Las multi-series cumplen propiedades análogas a las series. Muchos de los criterios de convergencia para series pueden ser generalizados a multi-series. En particular, tenemos el siguiente muy útil criterio de convergencia.

Teorema 1 (Criterio de Comparación). Sean (a_Q) y (b_Q) dos multi-sucesiones en K tales que

- (1) $|a_Q| \leq |b_Q|$, para todo multi-índice Q .
- (2) $\sum_Q |b_Q|$ es convergente

Entonces $\sum_Q a_Q$ es absolutamente convergente.

Para los fines que nos hemos planteado, las definiciones y propiedades establecidas líneas arriba son suficientes. Para el lector interesado en tener más información sobre multi-series, recomendamos la referencia [1].

3. Series de Potencia de Varias Variables Complejas

Sea (a_Q) una multi-sucesión de números complejos y $z_0 \in \mathbb{C}^n$. Una *serie de potencias* de varias variables complejas, con centro en z_0 es una expresión del tipo

$$\sum_{Q \geq 0} a_Q (z - z_0)^Q$$

Observe que para $n = 1$ tenemos la definición usual de serie de potencias de una variable compleja.

Para simplificar la notación, de ahora en adelante, sólo consideramos series de potencias centradas en el origen.

Definición 1. Sea $S = \sum_{Q \geq 0} a_Q z^Q$ una serie de potencias de varias variables complejas. El dominio de convergencia de la serie S , denotado por D_S es el mayor subconjunto abierto de \mathbb{C}^n tal que la serie S es convergente, es decir si $U \subseteq \mathbb{C}^n$ es un abierto tal que S es convergente $\forall z \in U$, entonces $U \subseteq D_S$

Observaciones

1. $a \in D_S$ si y sólo si existe un polidisco $\Delta(a, r) \subseteq D_S$ tal que S es convergente $\forall z \in \Delta(a, r)$.
2. El siguiente resultado es una consecuencia inmediata de nuestras definiciones:

“Sea $S = \sum_{Q \geq 0} a_Q z^Q$ una serie de potencias convergente en el polidisco $\Delta(0, R)$, con $R = (R_1 \cdots, R_n)$. Si $0 < r_j < R_j \forall 1 \leq j \leq n$, entonces la serie converge absolutamente y uniformemente en el polidisco cerrado $\Delta[0, r]$, donde $r = (r_1, \cdots, r_n)$ ”.
Una consecuencia inmediata de esto es que $a \in D_S$ si y sólo si existe un $r \in (\mathbb{R}^+)^n$ poliradio (con coordenadas suficientemente pequeñas) tal que la serie S es absolutamente y uniformemente convergente en $\Delta[a; r] \subseteq D_S$.

3. Si $n = 1$ entonces D_S es un disco abierto centrado en el origen. En dimensión $n > 1$ ¿ D_S es un polidisco?

Para responder a la interrogante planteada en la observación 3, vamos a introducir algunos conceptos.

Definición 2. Sea $S = \sum_{Q \geq 0} a_Q z^Q$ una serie de potencias de varias variables complejas. Definimos el conjunto

$$\Gamma_S = \{r = (r_1, \dots, r_n) \in (\mathbb{R}_0^+)^n : \sum_{Q \geq 0} |a_Q| r^Q < \infty\}.$$

Observaciones

1. $\Gamma_S \subseteq (\mathbb{R}_0^+)^n$ donde $\mathbb{R}_0^+ = [0, +\infty[$.
2. $\Gamma_S \neq \emptyset$ puesto que $(0, 0, \dots, 0) \in \Gamma_S$.
3. Si $n = 1$ entonces $\Gamma_S = [0, R[$ ó $\Gamma_S = [0, R]$ donde R es el radio de

convergencia de la serie de potencias $S = \sum_{m=0}^{\infty} a_m z^m$.

Proposición 1. Si $r \in \Gamma_S$ y $0 \leq \rho \leq r$ (es decir $0 \leq \rho_j \leq r_j \forall 1 \leq j \leq n$, siendo $\rho = (\rho_1, \dots, \rho_n)$ y $r = (r_1, \dots, r_n)$) entonces $\rho \in \Gamma_S$.

Demostración

$0 \leq \rho^Q \leq r^Q$, $\forall Q$ multi-índice, luego $0 \leq |a_Q| \rho^Q \leq |a_Q| r^Q$ para todo Q multi-índice. Del Criterio de Comparación se sigue que

$$\sum_{Q \geq 0} |a_Q| \rho^Q \leq \sum_{Q \geq 0} |a_Q| r^Q < \infty.$$

En consecuencia $\rho \in \Gamma_S$.

Corolario. Sea $r = (r_1, \dots, r_n) \in \Gamma_S$. Si $z = (z_1, \dots, z_n) \in \mathbb{C}^n$ es tal que $|z_j| \leq r_j$, $\forall 1 \leq j \leq n$ entonces $\sum_{Q \geq 0} a_Q z^Q$ es absolutamente convergente.

Definición 3. Sea $S = \sum_{Q \geq 0} a_Q z^Q$. La base del dominio de convergencia de la serie S , denotada por B_S es el conjunto de todos los puntos $r \in (\mathbb{R}_0^+)^n$ (donde \mathbb{R}_0^+ denota a los números reales mayores o iguales que 0) que son interiores a Γ_S .

Observaciones

1. $r = (r_1, \dots, r_n) \in B_S$ si y sólo si existen $\delta_1, \dots, \delta_n > 0$ tales que si $|\rho_j - r_j| < \delta_j$ y $\rho_j \geq 0$ se tiene que $\rho = (\rho_1, \dots, \rho_n) \in \Gamma_S$.
2. B_S es un subconjunto abierto de $(\mathbb{R}_0^+)^n$.
3. B_S puede ser vacío. En efecto, es fácil construir una serie de potencias S tal que $\Gamma_S = (0, \dots, 0)$, por ejemplo $S = \sum_{Q \geq 0} Q! z^Q$.
4. Si $n = 1$ entonces $B_S = [0, R[$.

A continuación, vamos a caracterizar a los elementos r del dominio de convergencia S .

Proposición 2. $r \in B_S$ si y sólo si existe un $r' \in (\mathbb{R}_0^+)^n$ con $r' > r$ tal que $r' \in \Gamma_S$

Demostración

(\Rightarrow) Si $r = (r_1, \dots, r_n) \in B_S$, entonces existen $\delta_1, \dots, \delta_n > 0$ tales que

$$|\rho_j - r_j| < \delta_j \quad y \quad \rho_j \geq 0 \Rightarrow \rho = (\rho_1, \dots, \rho_n) \in \Gamma_S.$$

Tomando $r' = (r_1 + \frac{\delta_1}{2}, \dots, r_n + \frac{\delta_n}{2})$ tenemos que $r' > r$ y

$$r' \in (]r_1 - \delta_1, r_1 + \delta_1[\times \dots \times]r_n - \delta_n, r_n + \delta_n[) \cap (\mathbb{R}_0^+)^n$$

luego $r' \in \Gamma_S$.

(\Leftarrow) Supongamos que existe $r' = (r_1, \dots, r_n) \in (\mathbb{R}_0^+)^n$ con $r'_j > r_j (1 \leq j \leq n)$ tal que $r' \in \Gamma_S$.

Considerando $\delta_j = r'_j - r_j > 0 (1 \leq j \leq n)$, se tiene que si $\rho = (\rho_1, \dots, \rho_n) \in (]r_1 - \delta_1, r_1 + \delta_1[\times \dots \times]r_n - \delta_n, r_n + \delta_n[) \cap (\mathbb{R}_0^+)^n$ entonces $0 \leq \rho_j < r_j + \delta_j = r'_j \forall 1 \leq j \leq n$, es decir $0 \leq \rho \leq r'$. De la Proposición 1 se sigue que $\rho \in \Gamma_S$. De esta manera $r \in B_S$.

El siguiente teorema caracteriza a los elementos del dominio de convergencia en términos de su base.

Teorema 2. $a = (a_1, \dots, a_n) \in D_S$ si y sólo si $(|a_1|, \dots, |a_n|) \in B_S$.

Demostración

(\Rightarrow) Sea $a = (a_1, \dots, a_n) \in D_S$, por la Proposición 2, es suficiente hallar un $r' = (r'_1, \dots, r'_n) \in (\mathbb{R}_0^+)^n$ con $r'_j > |a_j| (1 \leq j \leq n)$ tal que $r' \in \Gamma_S$. Como

$a \in D_S$, existe un $r = (r_1, \dots, r_n) \in (\mathbb{R}_0^+)^n$ tal que $S = \sum_{Q \geq 0} a_Q z^Q$ es absolutamente convergente en $\Delta(a, r) = D_{r_1}(a_1) \times \dots \times D_{r_n}(a_n)$.

Dado $1 \leq j \leq n$, consideramos

$$\begin{aligned} r'_j &= \left| a_j + \frac{r_j}{2} \frac{a_j}{|a_j|} \right| = \left(1 + \frac{r_j}{2} \frac{1}{|a_j|} \right) |a_j| \\ &= \left(1 + \frac{r_j}{2} \frac{1}{|a_j|} \right) |a_j| = |a_j| + \frac{r_j}{2} > |a_j|. \end{aligned}$$

Luego $r' = (r'_1, \dots, r'_n) \in (\mathbb{R}_0^+)^n$ satisface $r'_j > |a_j|, \forall 1 \leq j \leq n$.

Por otro lado, si tomamos $b = (b_1, \dots, b_n)$ tal que

$$b_j = a_j + \frac{r_j}{2} \frac{a_j}{|a_j|} \quad (1 \leq j \leq n), \text{ se sigue que } |b_j - a_j| = \frac{r_j}{2} < r_j \text{ es decir } b \in \Delta(a, r).$$

Luego $\sum_{Q \geq 0} |a_Q| |b^Q| < \infty$.

Pero

$$|b^Q| = \left| \prod_{j=1}^n b_j^{q_j} \right| = \left| \prod_{j=1}^n \left(a_j + \frac{r_j}{2} \frac{a_j}{|a_j|} \right)^{q_j} \right| = \prod_{j=1}^n (r'_j)^{q_j} = (r')^Q,$$

es decir $\sum_{Q \geq 0} |a_Q| (r')^Q < \infty$. Se desprende que $r' \in \Gamma_S$.

(\Leftarrow) Si $(|a_1|, \dots, |a_n|) \in B_S$ entonces existe un $r' = (r'_1, \dots, r'_n) \in (\mathbb{R}_0^+)^n$ con $r'_j > |a_j| \forall 1 \leq j \leq n$ tal que $r' \in \Gamma_S$. Del Corolario de la Proposición 1 se sigue que si $z \in \Delta(0, r')$ entonces la serie de potencias $\sum_{Q \geq 0} a_Q z^Q$ es absolutamente convergente.

Tomando $\delta_j < r'_j - |a_j|$ ($1 \leq j \leq n$) y haciendo $\delta = (\delta_1, \dots, \delta_n) \in (\mathbb{R}^+)^n$ se sigue que $\Delta(a, \delta) \subseteq \Delta(0, r')$. De esta manera $\sum_{Q \geq 0} a_Q z^Q$ es absolutamente convergente en $\Delta(a, \delta)$. Hemos probado que $a \in D_S$.

Observación

En dimensión $n = 1$, el teorema anterior nos dice que

$$z \in D_S \Leftrightarrow |z| \in B_S = [0, R[\Leftrightarrow z \in D_R(0)$$

es decir, para serie de potencias de una variable su dominio de convergencia D_S es el disco $D_R(0)$.

4. Dominios de Reinhardt

Motivados en las propiedades que cumple el dominio de convergencia de una serie de potencias de varias variables, es que tenemos la siguiente

Definición 4. Sea $D \subseteq \mathbb{C}^n$ un conjunto abierto y denotemos por S^1 al círculo unitario de \mathbb{C} .

(1) Decimos que D es un Dominio de Reinhardt si y sólo si

$$(z_1, \dots, z_n) \in D \text{ y } \xi_1, \dots, \xi_n \in S^1 \Rightarrow (\xi_1 z_1, \dots, \xi_n z_n) \in D.$$

(2) Decimos que D es un Dominio de Reinhardt completo si y sólo si

$$(z_1, \dots, z_n) \in D \text{ y } \xi_1, \dots, \xi_n \in D_1(0) \Rightarrow (\xi_1 z_1, \dots, \xi_n z_n) \in D.$$

(3) La base del dominio de Reinhardt (completo o no) D es el conjunto B_D definido por

$$B_D = \{(|z_1|, \dots, |z_n|) : (z_1, \dots, z_n) \in D\} \subseteq (\mathbb{R}_0^+)^n.$$

Observaciones

1. En dimensión $n = 1$, tenemos que los discos $D_r(0)$ son dominios de Reinhardt completos con base el intervalo $[0, r[$, mientras que el anillo (abierto) $A(0; r_1, r_2)$ centrado en $0 \in \mathbb{C}$ y de radios inferior y superior r_1, r_2 respectivamente, el cual es definido como

$$A(0; r_1, r_2) = \{z \in \mathbb{C} : r_1 < |z| < r_2\};$$

es un dominio de Reinhardt (no completo) y su base es el intervalo $]r_1, r_2[$.

2. Es inmediato de la definición que todo dominio de Reinhardt $D \subseteq \mathbb{C}^n$ satisface la siguiente propiedad:
 $z = (z_1, \dots, z_n) \in D \Rightarrow S_{|z_1|}[0] \times \dots \times S_{|z_n|}[0] \subseteq D$
en donde $S_{|z_j|}[0]$ es el disco centrado en el origen y de radio $|z_j|$.
3. Análogamente, todo dominio de Reinhardt completo $D \subseteq \mathbb{C}^n$ satisface la siguiente propiedad:
 $z = (z_1, \dots, z_n) \in D \Rightarrow \Delta(0, |z_1|, \dots, |z_n|) \subseteq D$
4. La base B_D de todo dominio de Reinhardt $D \subseteq \mathbb{C}^n$ siempre está contenida en el dominio de Reinhardt. Más aún

$$B_D = D \cap (\mathbb{R}_0^+)^n$$
es decir B_D es un conjunto abierto relativo de $(\mathbb{R}_0^+)^n$
5. Es claro que la base de un dominio de Reinhardt no necesariamente es un abierto de \mathbb{R}^n .

Acabamos de ver que a todo dominio de Reinhardt $D \subseteq \mathbb{C}^n$ se le asocia su base B_D la cual es un abierto relativo de $(\mathbb{R}_0^+)^n$. A continuación, probaremos que lo recíproco también es cierto, es decir si nos damos un B subconjunto abierto relativo de $(\mathbb{R}_0^+)^n$, entonces existe un dominio de Reinhardt $D \subseteq \mathbb{C}^n$ tal que su base es justamente B .

Proposición 3. Sea B un subconjunto abierto relativo de $(\mathbb{R}_0^+)^n$, entonces

$$D = \{(\xi_1 r_1, \dots, \xi_n r_n) : (r_1, \dots, r_n) \in B, \xi_1, \dots, \xi_n \in S^1\} \subseteq \mathbb{C}^n$$

es un dominio de Reinhardt cuya base es B .

Demostración

Si $z = (z_1, \dots, z_n) \in D$ entonces por definición del conjunto D , existe $(r_1, \dots, r_n) \in B$ y existen $\xi_1, \dots, \xi_n \in S^1$ tal que $z_j = \xi_j r_j$, para todo $1 \leq j \leq n$.

Dados $\zeta_1, \dots, \zeta_n \in S^1$, se cumple

$$\begin{aligned} (\zeta_1 z_1, \dots, \zeta_n z_n) &= (\zeta_1(\xi_1 r_1), \dots, \zeta_n(\xi_n r_n)) \\ &= ((\zeta_1 \xi_1) r_1, \dots, (\zeta_n \xi_n) r_n) \in D \end{aligned}$$

luego D es un dominio de Reinhardt.

Por otro lado, sea $s = (s_1, \dots, s_n) \in B_D$, entonces existe $(z_1, \dots, z_n) \in D$ tal que $s_j = |z_j|$, $\forall 1 \leq j \leq n$. Pero, por definición del conjunto D , deben existir un $(r_1, \dots, r_n) \in B$ y $\xi_1, \dots, \xi_n \in S^1$ tales que $z_j = \xi_j r_j$, para todo $1 \leq j \leq n$, luego $s_j = |z_j| = |\xi_j r_j| = r_j$, es decir $s \in B$, luego $B_D \subseteq B$. El otro contenido es obvio.

Observaciones

1. La base B_D de un dominio de Reinhardt $D \subseteq \mathbb{C}^n$ determina completamente a D , por esta razón muchas propiedades de D pueden ser deducidas de las propiedades de su base.

2. Si $B_D \subseteq (\mathbb{R}_0^+)^n$ es la base de un dominio de Reinhardt completo $D \subseteq \mathbb{C}^n$, entonces se cumple

$$(r_1, \dots, r_n) \in B \Rightarrow [0, r_1] \times \dots \times [0, r_n] \subseteq B.$$

La demostración de esta propiedad es inmediata a partir de nuestras definiciones, por esta razón la omitimos.

3. Cuando $n = 2$ la propiedad de la observación anterior nos proporciona una condición geométrica muy sencilla para saber si un dominio de Reinhardt es completo. En efecto, si existe un $(r_1, r_2) \in B \subseteq (\mathbb{R}_0^+)^2$ tal que el rectángulo $[0, r_1] \times [0, r_2]$ no está contenido en B entonces B no puede ser base de un dominio de Reinhardt completo.

4. Ya hemos visto que los discos son ejemplos de dominios de Reinhard completos en \mathbb{C} mientras que los anillos son ejemplos de dominios de Reinhardt en \mathbb{C} . Recíprocamente todo dominio de Reinhard completo en \mathbb{C} es un disco mientras que los dominios de Reinhardt en \mathbb{C} son uniones numerables de anillos concéntricos, centrados en $0 \in \mathbb{C}$.

La prueba de estas afirmaciones son consecuencia inmediata de la propiedad establecida en la Observación 2.

Por esta razón, en los cursos de una variable compleja es innecesario considerar dominios de Reinhardt.

Para caracterizar los dominios de convergencia de una serie de potencia de varias variables, nos falta aún otro concepto.

Definición 5. Sea $B \subseteq (\mathbb{R}^+)^n$ un conjunto no vacío.

1. El Conjunto Logaritmo de B , denotado por $\log B$, se define como $\log B = \{(\log r_1, \dots, \log r_n) : (r_1, \dots, r_n) \in B\}$.
2. Decimos que B es Logarítmicamente Convexo si y sólo si $\log B$ es un subconjunto convexo de \mathbb{R}^n .

Observaciones

1. Si $B =]r_1, r_2[\mathbb{R}^+$ entonces $\log B =]\log r_1, \log r_n[\subseteq \mathbb{R}$. Se deduce que los intervalos contenidos en el semieje positivo son conjuntos logarítmicamente convexos.
2. $\log B$ es un conjunto homeomorfo a B .
3. No es difícil probar que $B \subseteq (\mathbb{R}^+)^n$ es un conjunto logarítmicamente convexo si y sólo si se cumple la siguiente propiedad

$$(r_1, \dots, r_n), (s_1, \dots, s_n) \in B \Rightarrow (r_1^{(1-\alpha)} s_1^\alpha, \dots, r_n^{(1-\alpha)} s_n^\alpha) \in B, \forall \alpha \in [0,1]$$

Estamos ahora listos para dar la caracterización prometida.

Teorema 3. Si $S = \sum_{Q \geq 0} a_Q z^Q$ es una serie de potencias de varias variables entonces su dominio de convergencia D_S es un dominio de Reinhardt completo con base B_S logarítmicamente convexa.

Demostración

Si $(z_1, \dots, z_n) \in D_S$ entonces por el Teorema 2 se tiene que $(|z_1|, \dots, |z_n|) \in B_S$.

De la Proposición 2 se sigue que existe un $r = (r_1, \dots, r_n) \in \Gamma_S$ tal que $r_j > |z_j|$, $(1 \leq j \leq n)$.

Sean $\xi_1, \dots, \xi_n \in D_1[0]$ entonces $|\xi_j z_j| \leq |z_j| < r_j$, para todo $1 \leq j \leq n$. De esta manera $(|\xi_1 z_1|, \dots, |\xi_n z_n|) \in B_S$ y usando nuevamente la Proposición 2 tenemos que $(\xi_1 z_1, \dots, \xi_n z_n) \in D_S$. Esto prueba que D_S es un dominio de Reinhardt completo.

Por otro lado, si $(r_1, \dots, r_n), (s_1, \dots, s_n) \in B_S$ entonces existe un $t = (t_1, \dots, t_n) \in \Gamma_S$ tal que $t_j > r_j$ y $t_j > s_j$, para todo $1 \leq j \leq n$. Pero para cualquier $1 \leq j \leq n$ se tiene

$$r_j^{1/2} s_j^{1/2} = (r_j s_j)^{1/2} \leq \frac{r_j + s_j}{2} < t_j$$

esto implica que $(\sqrt{r_1 s_1}, \dots, \sqrt{r_n s_n}) \in B_S$. Así, hemos probado que B_S es logarítmicamente convexo.

Bibliografía

- [1] H. Cartan, *Teoría elemental de funciones analíticas de una o varias variables complejas*, Selecciones Científicas, 1968.
- [2] J. Conway, *Functions of one complex variable*, Vol. I, Springer-Verlag, 1991.
- [3] R. Gunning, *Analytic Functions of several complex variable*, Vol. I: *Function Theory*, Wadsworth & Brooks/Cole, 1990.
- [4] R. Gunning - H. Rossi, *Analytic functions of several complex variables*, Prentice-Hall, 1965.
- [5] L. Hormander, *An introduction to complex analysis in several variables*, D. Van Nostrand Company, 1966.

Gradiente Generalizado

*Joseph Wilmer Simpe Laura**
*Mg. William Echegaray Castillo***

RESUMEN

Aquí el estudio del Gradiente Generalizado es dirigido para condiciones no diferenciables, para esto, se hace un adecuado desarrollo del Gradiente Generalizado en sociedad con otras herramientas, obteniéndose así el propósito deseado, este estudio se realiza en general en un espacio de Banach. También se presenta una Aplicación, ejemplos y gráficos ilustrativos para su mejor entendimiento.

* Universidad Nacional de Ingeniería. Facultad de Ciencias.

** Asesor.

Introducción

En el presente trabajo se desarrollará la Teoría y el Cálculo del Gradiente Generalizado.

A partir de la Derivada Direccional y del Gradiente, que son términos ya conocidos, se pretende generalizar los mismos bajo ciertas condiciones. Comenzando con el caso de una función Localmente Lipschitz de valor real definida en un espacio de Banach. Luego cuando la función es continuamente diferenciable el Gradiente Generalizado será la derivada, o cuando la función es convexa se le asociará al subdiferencial del análisis convexo.

Se establece una serie de Cálculos Básicos que serán importantes para el desarrollo de los siguientes capítulos.

También se desarrollará una Teoría Geométrica Asociada de Conos Tangentes y Normales, y veremos la relación entre estos conceptos y su contraparte en el caso no diferenciable y en análisis convexo. Además trataremos una definición extendida del Gradiente Generalizado a partir de funciones Lipschitz o No Lipschitz.

Finalmente, ilustraremos el tema mediante una Aplicación en Optimización y no Diferenciabilidad (siendo el gradiente generalizado una alternativa para condiciones no diferenciables).

1. Conceptos preliminares

Consideremos el Espacio de Banach X cuyos elementos x serán vectores o puntos.

1.1 La condición Lipschitz

Definición 1.1.1 Sea $Y \subset X$ un subconjunto de X . Una función $f: Y \rightarrow \mathbb{R}$ se dice que satisface la condición Lipschitz (en Y), si para algún $K \geq 0$, se tiene

$$|f(y) - f(z)| \leq K\|y - z\| \quad \forall y, z \in Y$$

ésta es referida como una condición Lipschitz de rango K .

Definición 1.1.2 Se dice que f es Lipschitz (de rango K) próximo a x si, para algún $\varepsilon > 0$, f satisface la condición Lipschitz (de rango K) en el conjunto $x + \varepsilon B$ (es decir, dentro de una ε -vecindad de x).

1.2 La derivada direccional generalizada

Definición 1.2.1 Sea f Lipschitz próximo a x , sea $v \in X$ cualquier otro vector. La Derivada Direccional Generalizada de f en x en la dirección v , denotada por $f^\circ(x; v)$, es definida de la siguiente manera:

$$f^\circ(x; v) = \limsup_{\substack{y \rightarrow x \\ t \downarrow 0}} \frac{f(y + tv) - f(y)}{t},$$

donde $y \in X$, $t > 0$.

1.3 El Gradiente Generalizado

Definición 1.3.1 El Gradiente Generalizado de f en x , denotado por $\partial f(x) \subset X^*$, es un subconjunto de X^* , dado por

$$\partial f(x) = \{\zeta \in X^* / f^\circ(x; v) \geq \langle \zeta, v \rangle, \forall v \in X\},$$

Denotaremos por $\|\zeta\|_*$ la Norma en X^* :

$$\|\zeta\|_* := \sup\{\langle \zeta, v \rangle : \forall v \in X, \|v\| \leq 1\},$$

y B_* denota la bola unitaria abierta en X^* .

Proposición 1.3.1 Sea f Lipschitz de rango K próximo a x , entonces:

(a) $\partial f(x) \subset X^*$ es un subconjunto no vacío, convexo, débil* compacto de X^* y $\|\zeta\|_* \leq K$, $\forall \zeta \in \partial f(x)$.

(b) $\forall v \in X$, se tiene $f^\circ(x; v) = \max\{\langle \zeta, v \rangle : \zeta \in \partial f(x)\}$.

Ejemplo:

Calcularemos el gradiente generalizado de la función $f(x) = |x|$ en $X = \mathbb{R}$, la cual es Lipschitz por la desigualdad triangular ($|f(x) - f(y)| = ||x| - |y|| \leq |x - y|$).

Si $x > 0$

$$f^\circ(x;v) = \lim_{\substack{y \rightarrow x \\ t \downarrow 0}} \sup \frac{|y+tv| - |y|}{t} = \begin{cases} v & \text{si } v \geq 0 \\ v & \text{si } v < 0 \wedge y+tv \geq 0 \end{cases}$$

$$\Rightarrow f^\circ(x;v) = v \Rightarrow \partial f(x) = \{\zeta / v \geq \langle \zeta, v \rangle \forall v\} \Rightarrow \partial f(x) = \{1\}.$$

Si $x < 0$

$$f^\circ(x;v) = \lim_{\substack{y \rightarrow x \\ t \downarrow 0}} \sup \frac{|y+tv| - |y|}{t} = \begin{cases} -v & \text{si } v \geq 0 \wedge y+tv < 0 \\ -v & \text{si } v < 0 \end{cases}$$

$$\Rightarrow f^\circ(x;v) = -v \Rightarrow \partial f(x) = \{\zeta / (-v) \geq \langle \zeta, v \rangle, \forall v\} \Rightarrow \partial f(x) = \{-1\}.$$

Si $x = 0$

$$f^\circ(0;v) = \lim_{\substack{y \rightarrow 0 \\ t \downarrow 0}} \sup \frac{|y+tv| - |y|}{t} = \begin{cases} v & \text{si } v \geq 0 \\ -v & \text{si } v < 0 \end{cases}$$

$$\Rightarrow f^\circ(0;v) = |v| \Rightarrow \partial f(0) = \{\zeta / |v| \geq \langle \zeta, v \rangle, \forall v\} \Rightarrow \partial f(0) = [-1, 1].$$

1.4 Funciones Soporte

Se nota claramente que la Proposición 1.3.1 es equivalente a conocer $\partial f(x)$ o la función $f^\circ(x; \cdot)$, cada una es obtenida de la otra. Ahora caracterizaremos los conjuntos convexos cerrados por sus funciones soporte.

Definición 1.4.1 Sea $\Sigma \subset X^*$ un subconjunto de X^* , su función Soporte es definida en X^{**} , si $X \subset X^{**}$ un subconjunto de X^{**} , entonces $\forall x \in X$ se tiene:

$$\sigma_\Sigma(x) := \sup\{\langle \zeta, x \rangle : \zeta \in \Sigma\}.$$

Proposición 1.4.1 Sean $C, D \subset X$ subconjuntos no vacíos, cerrados, convexos de X , y sean $\Sigma, \Delta \subset X^*$ subconjuntos no vacíos, débil*-cerrados, convexos de X^* , entonces:

$$\text{a) } C \subset D \Leftrightarrow \sigma_C(\zeta) \leq \sigma_D(\zeta), \forall \zeta \in X^*.$$

$$\text{b) } \Sigma \subset \Delta \Leftrightarrow \sigma_\Sigma(x) \leq \sigma_\Delta(x), \forall x \in X.$$

1.5 Derivadas y Subderivadas

El principal resultado en esta Sección es que ∂f se reduce a la derivada si f es continuamente diferenciable, y al subdiferencial del análisis convexo si f es convexo.

1.5.1 Derivada Clásica

Definición 1.5.1. Sea $F : X \rightarrow Y$ una función de X hacia otro espacio de Banach Y . La usual Derivada Direccional (unilateral) de la F en x en la dirección v es

$$F'(x;v) := \lim_{t \downarrow 0} \frac{F(x+tv) - F(x)}{t}$$

Definición 1.5.2. Se dice que F admite una derivada de Gâteaux en x , un elemento en el espacio $\mathcal{L}(X, Y)$ de funcionales lineales continuas de X hacia Y denotado por $DF(x)$, si $\exists F'(x;v), \forall v \in X$ y $F'(x;v) = \langle DF(x), v \rangle$

1.5.2 Diferenciabilidad Estricta

Definición 1.5.3. Se dice que F admite una Derivada Estricta en x , un elemento de $\mathcal{L}(X, Y)$ denotado por $D_s F(x)$, tal que $\forall v$ se tiene

$$\lim_{\substack{y \rightarrow x \\ t \downarrow 0}} \frac{F(y+tv) - F(y)}{t} = \langle D_s F(x), v \rangle,$$

y que la convergencia es uniforme con respecto a v en conjuntos compactos (la cual es verdadera, si F es Lipschitz próximo a x).

Proposición 1.5.1. Sea F una función en una vecindad de x hacia Y , y sea ζ un elemento de $\mathcal{L}(X, Y)$, las siguientes afirmaciones son equivalentes:

- (a) F es estrictamente diferenciable en x y $D_s F(x) = \zeta$
- (b) F es Lipschitz próximo a x , y $\forall v \in X$ se tiene

$$\lim_{\substack{y \rightarrow x \\ t \downarrow 0}} \frac{F(y+tv) - F(y)}{t} = \langle \zeta, v \rangle.$$

Corolario 1.5.1. Supongamos que F es continuamente diferenciable en x , entonces F es estrictamente diferenciable en x y Lipschitz próximo a x .

Proposición 1.5.2 Sea f Lipschitz próximo a x y admite una derivada de Gâteaux $Df(x)$ (Hadamard o Estricta), entonces $Df(x) \in \partial f(x)$.

Ejemplo:

Consideremos la función

$$f(x) = \begin{cases} x^2 \operatorname{sen}\left(\frac{1}{x}\right) & \text{si } x \neq 0 \\ 0 & \text{si } x = 0 \end{cases}$$

Probaremos que f es Lipschitz próximo a 0, calcularemos $\partial f(0)$ y verificaremos que $Df(0) \in \partial f(0)$.

Veamos que f es Lipschitz próximo a 0.

como $f'(0) = \lim_{x \rightarrow 0} \frac{f(x) - f(0)}{x - 0} = \lim_{x \rightarrow 0} x \operatorname{sen} \frac{1}{x} = 0 \Rightarrow f$ es diferenciable en 0, y además $f'(0)$ se nota que es acotada en una vecindad de 0
 $\Rightarrow f$ es Lipschitz próximo a 0.

Calculando $\partial f(0)$

Como $f^\circ(0; v) = \limsup_{\substack{y \rightarrow 0 \\ t \downarrow 0}} \frac{f(y+tv) - f(y)}{t}$

$$\Rightarrow f^\circ(0;v) = \limsup_{\substack{y \rightarrow 0 \\ t \downarrow 0}} \frac{(y+tv)^2 \operatorname{sen}\left(\frac{1}{y+tv}\right) - y^2 \operatorname{sen}\left(\frac{1}{y}\right)}{t},$$

aplicando el Teorema de L'hospital respecto de t , tenemos

$$\Rightarrow f^\circ(0;v) = \limsup_{\substack{y \rightarrow 0 \\ t \downarrow 0}} 2v(y+tv) \operatorname{sen}\left(\frac{1}{y+tv}\right) + (y+tv)^2 \cos\left(\frac{1}{y+tv}\right) \frac{-v}{(y+tv)^2}$$

$$\Rightarrow f^\circ(0;v) = \limsup_{\substack{y \rightarrow 0 \\ t \downarrow 0}} 2v(y+tv) \operatorname{sen}\left(\frac{1}{y+tv}\right) + \cos\left(\frac{1}{y+tv}\right) (-v),$$

ahora hacemos lo siguiente, sea

$$g = 2v(y+tv) \operatorname{sen}\left(\frac{1}{y+tv}\right) \quad \text{y} \quad h = -v \cos\left(\frac{1}{y+tv}\right),$$

por propiedad de límites se sabe que

$$\limsup h + \liminf g \leq \limsup (h + g) \leq \limsup h + \limsup g,$$

para nuestro caso $\liminf_{\substack{y \rightarrow 0 \\ t \downarrow 0}} g = \limsup_{\substack{y \rightarrow 0 \\ t \downarrow 0}} g = 0$,

además $-1 \leq \cos\left(\frac{1}{y+tv}\right) \leq 1 \wedge -1 \leq -\cos\left(\frac{1}{y+tv}\right) \leq 1$, luego

$$\text{Si } v \geq 0 \Rightarrow \limsup_{\substack{y \rightarrow 0 \\ t \downarrow 0}} -v \cos\left(\frac{1}{y+tv}\right) = v \limsup_{\substack{y \rightarrow 0 \\ t \downarrow 0}} -\cos\left(\frac{1}{y+tv}\right) = v, \text{ y}$$

$$\text{Si } v < 0 \Rightarrow \limsup_{\substack{y \rightarrow 0 \\ t \downarrow 0}} -v \cos\left(\frac{1}{y+tv}\right) = -v \limsup_{\substack{y \rightarrow 0 \\ t \downarrow 0}} -\cos\left(\frac{1}{y+tv}\right) = -v,$$

luego por el Teorema del Sandwich

$$\Rightarrow f^\circ(0;v) = \limsup_{\substack{y \rightarrow 0 \\ t \downarrow 0}} (h + g) = \begin{cases} v & \text{si } v \geq 0 \\ -v & \text{si } v < 0 \end{cases} \Rightarrow f^\circ(0;v) = |v|,$$

$$\Rightarrow \partial f(0) = \{\zeta / |v| \geq \zeta v, \forall v\} \Rightarrow \partial f(0) = [-1, 1].$$

Calculando $Df(0)$.

$$\text{Como } \langle Df(x), v \rangle = \lim_{t \downarrow 0} \frac{f(x+tv) - f(x)}{t}$$

$$\Rightarrow \langle Df(0), v \rangle = \lim_{t \downarrow 0} \frac{f(tv) - f(0)}{t} = \lim_{t \downarrow 0} \frac{t^2 v^2 \sin\left(\frac{1}{tv}\right)}{t} = \lim_{t \downarrow 0} tv^2 \sin\left(\frac{1}{tv}\right) = 0$$

$$\Rightarrow \langle Df(0), v \rangle = 0 \quad \forall v \in \mathbb{R} \Rightarrow Df(0) = 0.$$

$$\Rightarrow Df(0) = 0 \in [-1, 1] = \partial f(0).$$

Proposición 1.5.3 Si f es estrictamente diferenciable en x , entonces f es Lipschitz próximo a x y $\partial f(x) = \{D_s f(x)\}$.

Proposición 1.5.4 Si f es Lipschitz próximo a x y $\partial f(x) = \{\zeta\}$, entonces f es estrictamente diferenciable en x y $D_s f(x) = \zeta$.

1.5.3 Subdiferencial

Definición 1.5.4 Sea $U \subset X$ un subconjunto convexo de X . Como sabemos, una función $f : U \rightarrow \mathbb{R}$ se dice que es convexa, si $\forall x, y \in U, \forall \lambda \in [0, 1]$ se tiene

$$f(\lambda x + (1 - \lambda)y) \leq \lambda f(x) + (1 - \lambda)f(y).$$

Definición 1.5.5 Sea $U \subset X$ un subconjunto de X . Sea $f : U \rightarrow \mathbb{R}$ una función (convexa), el Subdiferencial de f en x es definido por el conjunto de $\zeta \in X^*$ que satisfacen

$$f(y) - f(x) \geq \langle \zeta, y - x \rangle, \forall y \in U.$$

Proposición 1.5.5 Cuando f es convexa en U y Lipschitz próximo a x , entonces $\partial f(x)$ coincide con el subdiferencial en x , en el sentido del análisis convexo, y $f^\circ(x; v)$ coincide con la derivada direccional $f'(x; v), \forall v$.

Ejemplo:

Determinar el gradiente generalizado de la función $f : \mathbb{R}^n \rightarrow \mathbb{R}$ definida por

$$f(x_1, x_2, \dots, x_n) = \max \{x_i / i = 1, 2, \dots, n\}.$$

Vemos que f es Lipschitz próximo a (x_1, x_2, \dots, x_n) .

Veamos que f es convexa

Sean $(x_1, x_2, \dots, x_n), (y_1, y_2, \dots, y_n) \in \mathbb{R}^n$ y $\lambda \in [0, 1]$, así

$$\begin{aligned} f(\lambda(x_1, x_2, \dots, x_n) + (1 - \lambda)(y_1, y_2, \dots, y_n)) &= \\ \max \{\lambda x_i + (1 - \lambda) y_i / i = 1, 2, \dots, n\} &\leq \\ \max \{\lambda x_i / i = 1, 2, \dots, n\} + \max \{(1 - \lambda) y_i / i = 1, 2, \dots, n\} &= \\ \lambda \max \{x_i / i = 1, 2, \dots, n\} + (1 - \lambda) \max \{y_i / i = 1, 2, \dots, n\} &= \\ \lambda f(x_1, x_2, \dots, x_n) + (1 - \lambda) f(y_1, y_2, \dots, y_n) &\Rightarrow f \text{ es convexa.} \end{aligned}$$

Sea $I(x) = \{i / f_i(x) = f(x)\}$ los índices donde el máximo f es obtenido.

Calcularemos $f'(x; v)$.

$$f'(x; v) = \lim_{t \downarrow 0} \frac{\max_i \{x_i + tv_i\} - \max_i \{x_i\}}{t}$$

aplicando el Teorema de L'hospital respecto de t , se tiene

$$\Rightarrow f'(x; v) = \lim_{t \downarrow 0} \max_{i \in I(x)} \{v_i\} = \max_{i \in I(x)} v_i,$$

además como f es Lipschitz y convexa

$$\Rightarrow \text{por la Proposición 1.5.5 } f^\circ(x; v) = f'(x; v).$$

$$\Rightarrow \partial f(x) = \{\zeta \in \mathbb{R}^n / \max_{i \in I(x)} v_i \geq \langle \zeta, v \rangle, \forall v \in \mathbb{R}^n\}.$$

Veamos la ley de formación del $\partial f(x)$.

Si $n=1$ tenemos:

$$\max_{i \in I(x)} v_i = v^* \geq \langle \zeta, v \rangle, \forall v \in \mathbb{R} \Rightarrow v^* \geq \langle \zeta, v^* \rangle \Rightarrow \zeta = 1.$$

Si $n=2$ tenemos:

$$\max_{i \in I(x)} v_i \geq \langle \zeta, v \rangle = \sum_{i=1}^2 \zeta_i v_i \quad \text{donde } \zeta = (\zeta_1, \zeta_2), v = (v_1, v_2)$$

Si $i = 1, 2 \in I(x)$.

Supongamos que $v_1 \geq v_2$ y $\zeta_1 \geq 0$

$$\begin{aligned} \Rightarrow \max_{i \in I(x)} v_i = v_1 &\geq \zeta_1 v_1 + \zeta_2 v_2 \geq \zeta_1 v_2 + \zeta_2 v_2 \\ &= (\zeta_1 + \zeta_2) v_2 \Rightarrow v_1 \geq (\zeta_1 + \zeta_2) v_2. \end{aligned}$$

\Rightarrow tenemos dos casos:

Si $v_2 \geq (\zeta_1 + \zeta_2) v_2 \forall v_2 \Rightarrow \zeta_1 + \zeta_2 = 1$,

Si $(\zeta_1 + \zeta_2) v_2 \geq v_2, \forall v_2 \Rightarrow \zeta_1 + \zeta_2 = 1$

Continuamos con $v_1 \geq v_2$

$$\begin{aligned} \Rightarrow \max_{i \in I(x)} v_i = v_1 &\geq \zeta_1 v_1 + \zeta_2 v_2 = \zeta_1 v_1 + v_2 - \zeta_1 v_2 \\ &\Rightarrow v_1 - v_2 \geq \zeta_1 (v_1 - v_2) \Rightarrow 1 \geq \zeta_1, \end{aligned}$$

$$\begin{aligned} \text{además } v_1 &\geq \zeta_1 v_1 + \zeta_2 v_2 = v_1 - \zeta_2 v_2 \Rightarrow \\ &\Rightarrow (v_1 - v_2) \zeta_2 \geq 0 \Rightarrow \zeta_2 \geq 0 \Rightarrow \zeta_2 \leq 1. \end{aligned}$$

$$\Rightarrow \zeta = (\zeta_1, \zeta_2) / \zeta_1 + \zeta_2 = 1, \zeta_i \geq 0 \quad \text{donde } i = 1, 2 \in I(x)$$

además $\zeta_i = 0$ si $i \notin I(x)$.

Finalmente el resultado general será:

$$\partial f(x) = \left\{ (\zeta_1, \zeta_2, \dots, \zeta_n) / \zeta_i \geq 0, \sum_{i \in I(x)} \zeta_i = 1; \zeta_i = 0 \text{ si } i \notin I(x) \right\}.$$

2. Cálculos básicos

2.1 Relaciones fundamentales

Proposición 2.1.1 (Múltiplo escalar)

$\forall s \in \mathbb{R}$, se tiene

$$\partial(sf)(x) = s\partial f(x).$$

Proposición 2.1.2 (Extremo local)

Si f posee un mínimo local o un máximo local en x , entonces $0 \in \partial f(x)$.

Proposición 2.1.3 (Suma finita)

$$\partial\left(\sum_{i=1}^n f_i\right)(x) \subset \sum_{i=1}^n \partial f_i(x).$$

Corolario 2.1.1 En la Proposición 2.1.3 tenemos la igualdad si todas las funciones f_i son estrictamente diferenciables en x .

Corolario 2.1.2 $\forall s_i \in \mathbb{R}$ se tiene

$$\partial\left(\sum_{i=1}^n s_i f_i\right)(x) \subset \sum_{i=1}^n s_i \partial f_i(x)$$

Tenemos la igualdad si todas las funciones f_i son estrictamente diferenciables en x .

Observación:

En la Proposición 2.1.3 tenemos la igualdad si cada función f_i es continuamente diferenciable en x .

Por el Corolario 1.5.1 \Rightarrow cada función f_i es estrictamente diferenciable en x , luego se procede como en la demostración del Corolario 2.1.1

\Rightarrow Se obtiene la igualdad en la Proposición 2.1.3

Además de la Proposición 1.5.3 $\partial f_i(x) = \{D_s f_i(x)\} = \{D f_i(x)\}$, la Derivada.

2.2 Regularidad de Funciones

En el cálculo de fórmulas del gradiente generalizado a veces encontramos inclusiones como la de la Proposición 2.1.3, podemos agregar hipótesis para llevar cada regla de inclusión a igualdad. Como en la observación anterior, se nota que la condición continuamente diferenciable es fuerte, pues el gradiente generalizado es la derivada. Se desea una condición menos extrema, una función que cubra el caso no diferenciable, una condición útil es la siguiente:

Definición 2.2.1 f es Regular en x si:

- i) Existe la derivada direccional usual unilateral $f'(x;v) \forall v$.
- ii) $\forall v \ f'(x;v) = f^\circ(x;v)$.

Corolario 2.2.1 Si cada f_i es regular en x , entonces tenemos la igualdad en la Proposición 2.1.3. También $\forall s_i \geq 0$ tenemos la igualdad en el Corolario 2.1.2.

2.3 Teorema del Valor Medio

Teorema 2.3.1 (Lebourg). Sea $x, y \in X$, y supongamos que f es Lipschitz en un conjunto abierto conteniendo el segmento de línea $[x, y]$, entonces existe un punto $u \in \langle x, y \rangle$ tal que

$$f(y) - f(x) \in \langle \partial f(u), y - x \rangle$$

3. Conceptos Geométricos Asociados

Definimos una función distancia d_C no diferenciable, globalmente Lipschitz. Con el ∂d_C definimos un nuevo concepto de Tangentes y Normales en un conjunto C . Estas Tangentes y Normales no dependen de d_C . Veremos que las nuevas Tangentes y Normales, se reducen a conocer una de ellas en análisis convexo y en el caso no diferenciable.

3.1 La función distancia

Definición 3.1.1 Sea $C \subset X$ un subconjunto no vacío de X , y consideremos la función Distancia $d_C(.) : X \rightarrow \mathbb{R}$ definida por

$$d_C(x) = \inf \{ \|x - c\| : c \in C \}.$$

Proposición 3.1.1 La función d_C es global Lipschitz en X : $|d_C(x) - d_C(y)| \leq \|x - y\|$.

3.2 Tangentes

Definición 3.2.1 Supongamos que $x \in C$. Un vector $v \in X$ es tangente hacia C en x si $d_C^o(x; v) = 0$. El conjunto de todas las tangentes hacia C en x es denotado por $T_C(x)$ es decir

$$T_C(x) = \{v \in X / d_C^o(x; v) = 0\}.$$

Propiedades de $T_C(x)$: Cono convexo cerrado en X .

3.3 Normales

Definición 3.3.1 Sea M un cono convexo. El cono polar de M (Negativo) es:

$$M^\circ = \{s \in X^* / \langle s, v \rangle \leq 0, \forall v \in M\}.$$

Definición 3.3.2 El cono normal hacia C en x , es la Polaridad de $T_C(x)$:

$$N_C(x) = \{\zeta \in X^* \mid \langle \zeta, v \rangle \leq 0, \forall v \in T_C(x)\}.$$

Proposición 3.3.1

$$N_C(x) = cl \left\{ \bigcup_{\lambda=0} \lambda \partial d_C(x) \right\},$$

donde cl denota la clausura débil*.

Observación:

$\partial d_C(x)$ se reduce a $N_C(x)$ el cual es cerrado débil* por la Proposición 3.3.1. El cono polar hacia $T_C(x)$ es cerrado débil* ($N_C(x)$) cono convexo generado por $\partial d_C(x)$.

Propiedades de $N_C(x)$: Cono convexo en X^* .

Definición 3.3.3 Si C es convexo, se tiene un concepto bien definido de vector normal: $\zeta \in X^*$ se dice Normal hacia C en x , si $\langle \zeta, x - c \rangle \geq 0 \forall c \in C$. (en el sentido del análisis convexo).

Proposición 3.3.2 Si C es convexo, entonces $N_C(x)$ coincide con el cono de normales en el sentido del análisis convexo.

3.4 Caracterización Intrínseca de las Tangentes

Teorema 3.4.1 Un elemento $v \in X$ es tangente hacia C en x si, y sólo si, para cada sucesión $\{x_i\} \subset C$ convergente hacia x y la sucesión $\{t_i\} \subset \langle 0, \infty \rangle$ decreciente hacia cero, existe una sucesión $\{v_i\} \subset X$ que converge hacia v tal que $x_i + t_i v_i \in C$ para cada i .

3.5 Regularidad de Conjuntos

Se establece aquí la relación entre el concepto geométrico ya definido y las nociones en el caso no diferenciable, se requiere la regularidad de conjuntos, la cual utiliza la regularidad de funciones tratadas en el Capítulo 2.

Definición 3.5.1. Sea $K_C(x)$ el cono contingente de tangentes hacia un conjunto C en x . ($v \in X$). Un vector $v \in K_C(x)$ si, y sólo si, $\forall \varepsilon > 0, \exists t \in \langle 0, \varepsilon \rangle \wedge w \in v + \varepsilon B / x + tw \in C$ (donde w es un punto y necesariamente $x \in cl C$).

Observación: $T_C(x) \subset K_C(x)$

Definición 3.5.2 El conjunto C es regular en x si $T_C(x) = K_C(x)$.

Veamos ahora como T_C y N_C son obtenidos a partir de condiciones de regularidad.

Teorema 3.5.1 Sea f Lipschitz próximo a x , ahora supongamos que $0 \notin \partial f(x)$. Si $C = \{y \in X / f(y) \leq f(x)\}$, entonces se tiene

$$\{v \in X / f^\circ(x; v) \leq 0\} \subset T_C(x).$$

Además, si f es regular en x , entonces se tiene la igualdad en esta última inclusión, y también C es regular en x .

Corolario 3.5.1

$$N_C(x) \subset \bigcup_{\lambda \geq 0} \lambda \partial f(x).$$

Si f es regular en x , entonces se tiene la igualdad en esta última inclusión.

3.6 Epígrafos

Función distancia d_C nexa entre la teoría analítica del Gradiente Generalizado y la teoría Geométrica ya definida. Un vínculo diferente a través del Epígrafo.

Definición 3.6.1 El Epígrafo de una función $f: X \rightarrow \mathbb{R}$ es el siguiente subconjunto de $X \times \mathbb{R}$:

$$\text{epi}(f) := \{(x, r) \in X \times \mathbb{R} / f(x) \leq r\}.$$

Se observa que $\text{epi}(f)$ captura toda la información por sobre la función f .

Las siguientes relaciones se obtienen gracias al $\text{epi}(f)$.

Teorema 3.6.1 Sea f Lipschitz próximo a x , entonces

- (i) El Epígrafo de $f^\circ(x; \cdot)$ es $T_{\text{epi}(f)}(x, f(x))$, es decir $(v, r) \in T_{\text{epi}(f)}(x, f(x))$ si, y sólo si $r \geq f^\circ(x; v)$.
- (ii) f es regular en x si, y sólo si $\text{epi}(f)$ es regular en $(x, f(x))$.

Corolario 3.6.1 Sea $\zeta \in X^*$ es tal que $\zeta \in \partial f(x)$ si, y sólo si $(\zeta, -1) \in N_{\text{epi}(f)}(x, f(x))$.

Observación:

Sea $f: \mathbb{R} \rightarrow \mathbb{R}$ una función continuamente diferenciable en \mathbb{R} , según el Corolario anterior, el vector $(f'(x), -1)$ es normal hacia el grafo de la función f en $(x, f(x))$.

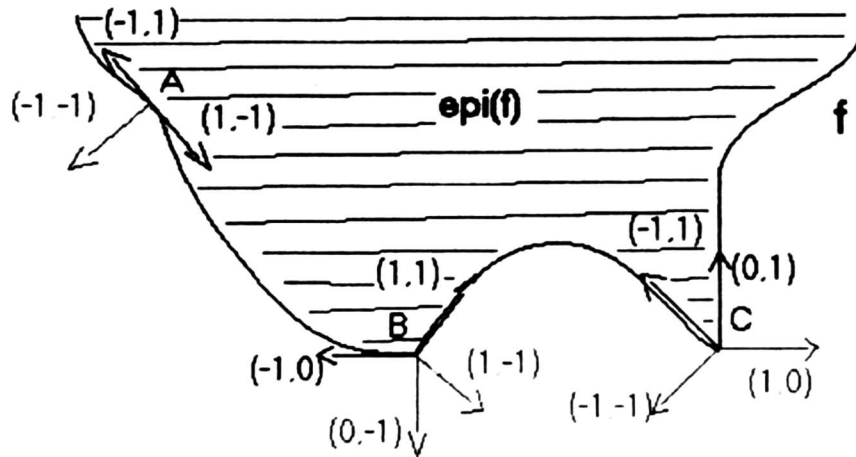
3.7 Una Extensión del Gradiente Generalizado con Funciones Lipschitz o No Lipschitz

Definición para una función f Lipschitz o No, del ∂f . El Corolario Anterior garantiza que esta nueva definición es consistente para f Lipschitz.

Definición 3.7.1 Sea $f : X \rightarrow \mathbb{R} \cup \{-\infty\} \cup \{+\infty\}$ es finito en x . Se define

$$\partial f(x) = \{\zeta \in X^* / (\zeta, -1) \in N_{\text{epi}(f)}(x, f(x))\}.$$

Ejemplo.- Calcular ∂f en los puntos indicados.



Nota: Los vectores de inclinación diagonal forman un ángulo de 45° con la horizontal.

En el punto A: f es Lipschitz en A , $\partial f = \{-1\}$ además f continuamente diferenciable.

En el punto B: f es Lipschitz en B , y además se nota que ζ tiene un recorrido de valores de $\zeta = 0$ hasta $\zeta = 1 \Rightarrow \partial f = [0, 1]$.

En el punto C: f no es continua, es decir no es Lipschitz en C , pero sin embargo todavía se puede usar la definición anterior para el caso lateral izquierdo, así $\partial f = [-1, +\infty)$, pues se tiene una pendiente vertical.

Aplicación

Ejemplo en Análisis no Diferenciable y Optimización, un ajuste de datos puntos.

Ejemplo

Considerando el problema de determinar la línea recta en el plano $X - Y$ que mejor ajuste los datos puntos. Para una recta dada $y = mx + b$ el error e_i en el i -ésimo dato punto (x_i, y_i) es definido por los valores de la recta aproximante y los valores dados $|mx_i + b - y_i|$. Una aproximación lineal requiere que la inclinación m y el intercepto b minimicen la función $\sum_{i=0}^N e_i$ no diferenciable.

Específicamente, para $N + 1$ datos puntos $(0, 0), (1, 1), \dots, (N - 1, N - 1)$ junto con $(N, 0)$, consideremos determinar la recta que mejor ajuste estos datos. Minimizar:

$$f(\alpha, \beta) = |\alpha N + \beta| + \sum_{i=1}^{N-1} |\alpha i + \beta - i|.$$

En donde la función $f_{c,k}(\alpha, \beta)$ es definida por $f_{c,k}(\alpha, \beta) = |\alpha c + \beta - k|$, ésta es la composición de g y F , donde $g(y) = |y|$, $F(\alpha, \beta) = \alpha c + \beta - k$.

Como f es no diferenciable, resolveremos el problema en base a lo previamente estudiado.

Calculando $D_s F(\alpha, \beta)$.

$$\langle D_s F(\alpha, \beta), v \rangle = \lim_{\substack{(\alpha', \beta') \rightarrow (\alpha, \beta) \\ t \downarrow 0}} \frac{F((\alpha', \beta') + t(v_1, v_2)) - F(\alpha', \beta')}{t} =$$

$$\lim_{\substack{(\alpha', \beta') \rightarrow (\alpha, \beta) \\ t \downarrow 0}} \frac{(\alpha' + tv_1)c + \beta' + tv_2 - k - \alpha'c - \beta' + k}{t} = v_1c + v_2 = \langle (c, 1), (v_1, v_2) \rangle$$

$\Rightarrow D_s F(\alpha, \beta) = (c, 1)$. $\Rightarrow F$ es estrictamente diferenciable en (α, β)

\Rightarrow por la Proposición 1.5.3 F es Lipschitz próximo a (α, β) . Y como g es Lipschitz

$\Rightarrow goF$ es Lipschitz próximo a (α, β) .

$\Rightarrow f_{c,k}(\alpha, \beta) = goF(\alpha, \beta) = |\alpha c + \beta - k|$ es Lipschitz próximo a (α, β) .

Calculando $\partial f_{c,k}(\alpha, \beta)$.

Si $\alpha c + \beta - k = 0$.

$$f_{c,k}^o((\alpha, \beta); v) = \lim_{\substack{(\alpha', \beta') \rightarrow (\alpha, \beta) \\ t \downarrow 0}} \frac{|(\alpha' + tv_1)c + (\beta' + tv_2) - k| - |\alpha'c + \beta' - k|}{t} \geq \langle \zeta, v \rangle$$

$$\Rightarrow \limsup_{\substack{(\alpha', \beta') \rightarrow (\alpha, \beta) \\ t \downarrow 0}} \frac{|\alpha'c + \beta' - k + t(v_1c + v_2)| - |\alpha'c + \beta' - k|}{t} = |v_1c + v_2| \geq \langle \zeta, v \rangle$$

$$\Rightarrow \partial f_{c,k}(\alpha, \beta) = \{\lambda(c, 1) / |\lambda| \leq 1\} \quad (*)$$

Sabemos que $f(\alpha, \beta) = |\alpha N + \beta| + \sum_{i=1}^{N-1} |\alpha i + \beta - i|$ y

$$f_{c,k}(\alpha, \beta) = |\alpha c + \beta - k|.$$

Para $\alpha c + \beta - k = 0$ (el análisis se realiza en esta recta), $\exists(\alpha, \beta)$ que minimiza $f \Rightarrow$ por la proposición 2.1.2 $\bar{0} \in \partial f(\alpha, \beta)$, donde $f = \sum_{i=0}^N f_{i,k}$ (sumatoria sólo respecto a i ; $k \in \mathbb{Z}_0^+$).

Además, por la proposición 2.1.3 $\partial \left(\sum_{i=0}^N f_{i,k} \right) (\alpha, \beta) \subset \sum_{i=0}^N \partial f_{i,k}(\alpha, \beta)$

$$\Rightarrow \bar{0} \in \sum_{i=0}^N \partial f_{i,k}(\alpha, \beta) = \sum_{i=0}^N \lambda_i(i, 1) \text{ ésto de } (*)$$

$$\Rightarrow \bar{0} = \lambda_N(N, 1) + \sum_{i=0}^{N-1} \lambda_i(i, 1) \quad (**)$$

f es convexa pues $f(\lambda(\alpha, \beta) + (1 - \lambda)(m, n)) \leq \lambda f(\alpha, \beta) + (1 - \lambda) f(m, n)$. Si $\bar{0} \in \partial f(\alpha, \beta)$ y como f es convexo

$\Rightarrow f(f, s) - f(\alpha, \beta) \geq \langle \bar{0}, (r, s) - (\alpha, \beta) \rangle \forall (r, s) \in U$ (subdiferencial del análisis convexo).

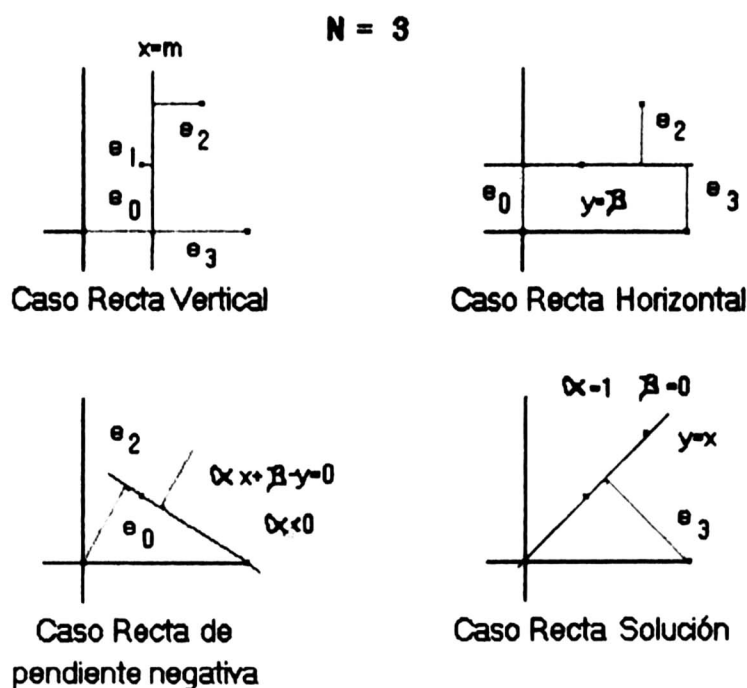
$\Rightarrow f(r, s) \geq f(\alpha, \beta) \forall (r, s) \in U \Rightarrow f$ alcanza un mínimo local en (α, β) .

Según nuestros datos puntos, mediante un análisis el punto $(\alpha = 1, \beta = 0)$ minimiza f .

$\Rightarrow k = c \Rightarrow$ la recta $y = x$ es la solución del problema.

Como un ejemplo, para $N = 3$, se tiene que:

- Se descartan las rectas verticales, pues pasan por a los más un punto y producen un error grande.
- Se descartan las rectas horizontales, pues pasan por a los más dos puntos y produce un error = 3.
- Se descartan las rectas de pendiente negativa, pues pasan por a los más dos puntos y producen un error grande.
- Nos damos cuenta que la recta adecuada es $y = x$ pues pasa por tres puntos y además produce un error = 3 (el más óptimo), las otras rectas producen errores grandes y sólo pasan por a los más un punto.



Se nota que $(**)$ es una condición necesaria para la recta $y = x$ ($\alpha = 1, \beta = 0$).

$$\Rightarrow \lambda_N = 1 \wedge |\lambda_i| \leq 1 \quad (i = 0, \dots, N - 1)$$

Si $N = 1$ tenemos $\lambda_1(1, 1) + \lambda_0(0, 1) = \bar{0}$, donde $\lambda_1 = 1 \Rightarrow 1 = 0$, así, no se satisface (**).

Si $N = 2$ tenemos $\lambda_2(2, 1) + \lambda_0(9, 1) + \lambda_1(1, 1) = \bar{0}$, donde $\lambda_2 = 1 \Rightarrow \lambda_1 = -2$, así, no se satisface (**).

Lo verificaremos para $N = 3$.

$$\Rightarrow \bar{0} = \lambda_3(3, 1) + \sum_{i=0}^2 \lambda_i(i, 1) = \lambda_3(3, 1) + \lambda_0(0, 1) + \lambda_1(1, 1) + \lambda_2(2, 1),$$

donde $\lambda_3 = 1 \Rightarrow 0 = 3 + \lambda_1 + 2\lambda_2 \wedge 0 = 1 + \lambda_0 + \lambda_1 + \lambda_2$;

para $\lambda_0 = 1, \lambda_1 = -1$ y $\lambda_2 = -1$ se verifican estas dos últimas ecuaciones.

$\Rightarrow \exists \lambda_i$ s que satisfacen (**). Si $N \geq 3 \Rightarrow \exists \lambda_i$ s que satisfacen (**).

Conclusiones

Algunos estudios en Análisis, que están envueltos sobre hipótesis de funciones continuamente diferenciables o diferenciables, pueden ser desarrollados también bajo condiciones no diferenciables, las cuales por ejemplo son utilizadas en la Optimización, tal y como ocurre en la Aplicación anterior.

El estudio del Gradiente Generalizado nos permite tener una alternativa de cálculo para con las funciones no diferenciables, pues sus distintas propiedades, cálculos básicos y asociaciones geométricas, nos da la posibilidad de afrontar problemas de optimización con estas alternativas.

En general podemos calcular el Gradiente Generalizado de una función. Cuando esta función es continuamente diferenciable su cálculo se simplifica, como se observa a lo largo de este trabajo en algunas observaciones. Por otro lado cuando se tiene una condición menos fuerte, (no diferenciable), ésta se calcula como se ha demostrado en sociedad con otras herramientas (Tangentes, Normales, Regularidad de Funciones y de Conjuntos, Epígrafos, etc.), tal y como se describe en el Capítulo 3. Y cuando la función es convexa se le asocia al subdiferencial de análisis convexo.

Un importante resultado es que el Epígrafo de una función nos permite una relación directa entre las Normales, Tangentes, Gradiente Generalizado y la Derivada Direccional Generalizada, pues conociendo una de éstas, se pueden obtener las demás. Gracias al Epígrafo, se pudo establecer una extensión de la definición del Gradiente Genralizado.

Bibliografía

- [1] Apostol Tom M., *Análisis Matemático*, Editorial Reverté, S.A., México, 1993.
- [2] Apostol Tom M., *Calculus*, Editorial Reverté, S.A., México, 1997.
- [3] Brézis Haim, *Análisis Funcional. Teoría y Aplicaciones*, Alianza Editorial, S.A., Madrid, 1984.
- [4] Clarke Frank H., *Optimization and Nonsmooth Analysis*, Editorial Board, USA, 1983.
- [5] Chumpitaz Mauro Reyna, *Teoría de la Medida*, Lima, Perú, 1987.
- [6] Fleitas Morales G., Margalef Roig J., Editorial Alhambra, S.A., Madrid, 1980.
- [7] Haaser N., Lasalle J., Sullivan J., *Análisis Matemático 2*, Editorial Trillas, México, 1972.
- [8] Lemaréchal Claude, *Convex Analysis and Minimization Algorithms I*, USA, 1993.
- [9] Purcell Edwin J., Varberg Dale, *Cálculo con Geometría Analítica*, Programas Educativos S.A. de C.V., México D.F., 1993.
- [10] Rockafellar R. Tyrrell, *Convex Analysis*, New Jersey, 1970.
- [11] Rudin Walter, *Análisis Funcional*, Editorial Reverté, S.A., Barcelona, 1979.
- [12] Takeuchi Yu, *Sucesiones y Series*, Editorial Limusa, México, 1980.

Estudio de la Composición Química de Multicapas Magnéticas en Función de la Profundidad

J. Carlos Rojas Sánchez¹ Hugo Ascolani² Laura Steren³

RESUMEN

El objetivo del trabajo es determinar la composición química de un arreglo de tri-capas magnéticas de $\text{La}_{0.55}\text{Sr}_{0.45}\text{MnO}_3/\text{SrTiO}_3/\text{La}_{0.55}\text{Sr}_{0.45}\text{MnO}_3$ en función de la profundidad, aplicando las técnicas de espectrometría de fotoelectrones y espectroscopía de electrones Auger. Se estudiaron dos muestras diferentes. En una de ellas la tri-capa magnética crecida sobre MgO había sido esculpida por métodos de litografía óptica para formar contactos eléctricos. En éste caso encontramos que el litografiado había resultado defectuoso por un error en la calibración de los tiempos de grabado iónico utilizados. En la segunda muestra se encontró que el método de sputtering no fue adecuado para investigar posibles efectos de interdifusión química en las interfases de estas muestras debido a que el tamaño de la región superficial dañada por el propio grabado iónico resultó demasiado grande comparada con los tamaños que interesan para el estudio.

¹ Facultad de Ciencias - Universidad Nacional de Ingeniería - crsunifis@hotmail.com.

² Director. ascolani@cab.cnea.gov.ar

³ Colaboradora. steren@cab.cnea.gov.ar Becas de verano 2003 Instituto Balseiro - Centro Atómico Bariloche.

Introducción

En los últimos diez años, se han producido enormes avances en la fabricación e investigación de nanoestructuras magnéticas artificiales. Se ha logrado confinar artificialmente materiales a bajas dimensiones – puntos, hilos y capas – dando lugar al descubrimiento de nuevos fenómenos físicos. El fenómeno de magnetoresistencia gigante, magnetorresistencia colosal, el desarrollo de **junturas túnel** y la inyección de espín son los resultados más asombrosos obtenidos en este campo de la Física donde se mezclan el magnetismo con el transporte electrónico. Un campo de rápido desarrollo en este área es el denominado «*espintrónica*», en el cual dos grados de libertad, la carga y el espín, están estrechamente ligados y son utilizados simultáneamente para crear nuevas funcionalidades. Típicamente, los espesores que son de interés para la espintrónica son de unos 400 \AA para las capas de manganita y de unos $10\text{-}50 \text{ \AA}$ para el separador. El objetivo del estudio es determinar la calidad de las interfases en cuanto a la posible existencia de difusión química o agujeros en la capa de titanato de estroncio (pin holes) que conecten las capas de manganita (ya no sería una juntura túnel). Para realizar este estudio se combinaron las técnicas de espectroscopía de fotoelectrones (XPS) y espectroscopía de electrones Auger, con la técnica de grabado iónico (sputtering).

Espectrosocopía de Electrones

Cuando se irradia un material con fotones, electrones o iones éste emite electrones en el rango de bajas energías (5-2000 eV) que contienen información sobre la composición química, la estructura electrónica, la estructura cristalina y las excitaciones de la muestra. Existe una variedad de técnicas experimentales basadas en la detección de los electrones emitidos diseñadas para extraer esta información, entre las cuales se encuentran la espectroscopía de fotoelectrones y la de electrones Auger [1]. En la figura 1 vemos el esquema del equipo usado en espectroscopía de electrones. La muestra se irradia con haces de fotones o electrones monocromáticos y se mide la distribución de energías cinéticas de los electrones emitidos, para lo cual se utilizan analizadores electrostáticos. Existen distintas geometrías para estos analizadores siendo las mas usadas la geometría cilíndrica, CMA por sus siglas en inglés (Cylindrical Mirror Analyser) o como es nuestro caso dos hemisferios concentricos, CHA

(Concentric hemispherical Analyser). El resultado de la medición es el número de cuentas (electrones) en función de la energía cinética

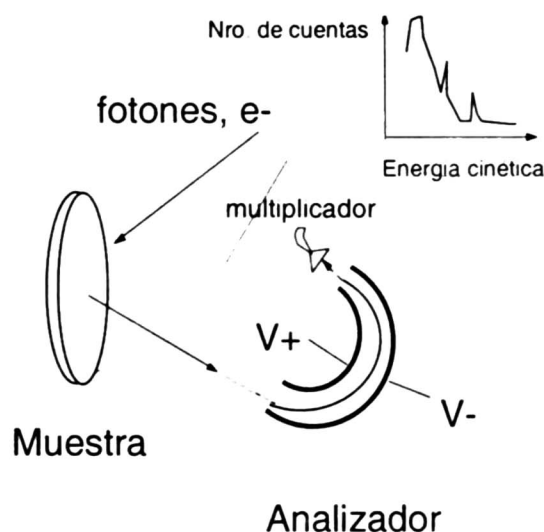


Figura 1. Esquema de un experimento de espectroscopía de electrones. La muestra es irradiada con fotones o electrones, emitiendo electrones característicos de la superficie, los electrones pasan a través de un analizador de energía para obtener un espectro de la distribución de electrones emitidos en función de su energía cinética.

Las técnicas de espectroscopía de electrones son útiles para el análisis superficial de la muestra porque, como puede verse en la Figura 2, los electrones con energías entre 5 – 2000 eV tienen un rango de penetración (longitud de atenuación) menor que 20 \AA . Un electrón pasando a través de un sólido puede perder energía de diferentes maneras que determinan la distancia recorrida dentro del sólido. La dispersión con fonones podemos despreciarla ya que produce cambios muy pequeños en la energía cinética. Los principales procesos inelásticos son la excitación de plasmones, la excitación de electrones de la banda de valencia, y la ionización de los niveles de las capas internas de los átomos que constituyen el sólido. El último proceso tiene una sección eficaz pequeña comparada con los otros procesos y su camino libre medio es generalmente dos órdenes de magnitud más grande que el camino libre medio de los otros dos procesos.

El grado de atenuación para electrones emitidos desde el substrato con una determinada energía que atraviesan una película de espesor d es igual a $\exp(-d/\lambda)$, donde λ es la longitud de atenuación.

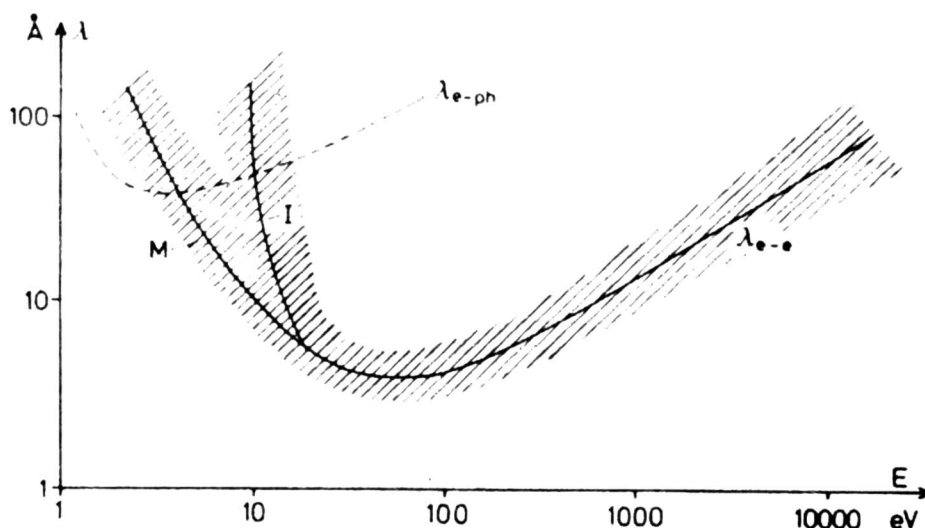


Figura 2. Dibujo esquemático de la dependencia de la longitud de atenuación de un electrón en un sólido en función de su energía. La mayoría de los valores experimentales caen dentro de la zona sombreada. (M = metal, I = aislante).

Los procesos de fotoemisión y de emisión Auger

En el proceso de fotoemisión un electrón de la muestra que ocupa un estado cuya energía de ligadura es E_A absorbe un fotón de energía $h\nu$ y pasa a un estado electrónico desocupado de energía

$$K = h\nu - E_A \quad (1)$$

Si la energía K es suficientemente grande el electrón excitado puede salir al vacío. El proceso de fotoemisión deja al sólido excitado con un hueco en un nivel interno.

La relajación del sistema se produce a través de dos fenómenos diferentes que compiten entre sí: la fluorescencia y la emisión de electrones Auger. Tanto en la fluorescencia como en la transición Auger el hueco en el nivel interno A es ocupado por un electrón que decae desde un nivel superior B. La diferencia entre ambos procesos es que en el primer caso la energía ($E_B - E_A$) disponible se emite como un fotón, mientras que en la transición Auger se emite un electrón de un nivel C. Por lo tanto, la energía K del electrón Auger que sale del nivel C será:

$$K = E_A - E_B - E_C \quad (2)$$

Es importante notar que, a diferencia de la fotoemisión donde la energía del electrón depende del fotón incidente, la energía de un electrón Auger sólo depende de diferencias entre las energías de los tres niveles involucrados. La transición Auger es más probable para elementos con bajo número atómico Z mientras que para altos Z el proceso de fluorescencia es el que predomina.

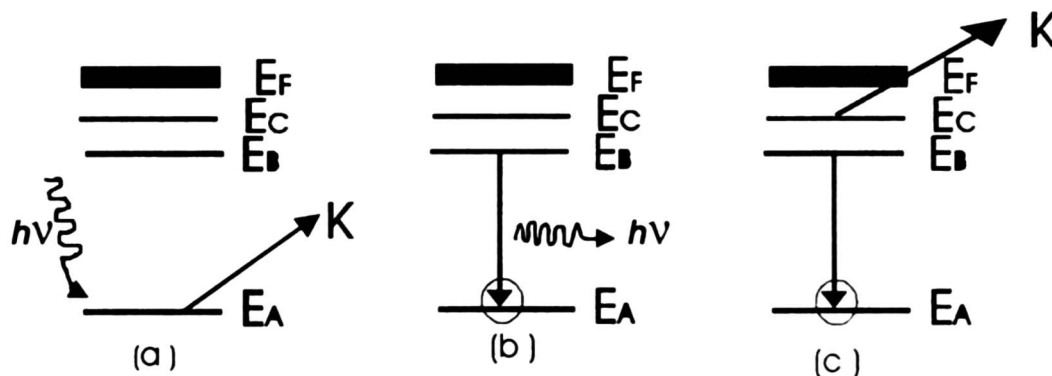


Figura 3. Procesos de (a) fotoemisión, (b) fluorescencia y (c) electrones Auger.

Espectroscopía de Fotoemisión de Rayos X (XPS)

La técnica de XPS se basa en el análisis de los electrones provenientes de niveles internos de los átomos superficiales de la muestra y para ello es necesario utilizar fotones con energías en el rango de los rayos X blandos. Como fuente de fotones se usa típicamente la línea $K\alpha$ de Mg (1253.6 eV) o Al (1486.6 eV) y más recientemente radiación de sincrotrón.

El resultado directo es el número de cuentas como función de la energía cinética de los electrones $N(K)$, la energía del haz de fotones con que se irradia la muestra es $h\nu$ y utilizando la ecuación (1) se puede obtener la energía de ligadura de los electrones E en vez de su energía cinética K . En un espectro XPS se presentan picos correspondientes a transiciones electrónicas de fotoemisión y también picos asociados a procesos Auger, se puede distinguir entre ellos midiendo el espectro con distintas energías de fotón ya que si se grafica las cuentas en función de la energía cinética, entonces los picos Auger de ambos espectros coinciden pero no los picos de fotoelectrones. Por el contrario, si se grafica las cuentas en función de la energía de ligadura los picos de fotoemisión coinciden pero no los correspondientes

a transiciones Auger. Este comportamiento se ilustra en la Figura 4 donde se muestra espectros XPS del compuesto LaFeO_3 producidos con los ánodos de magnesio y aluminio.

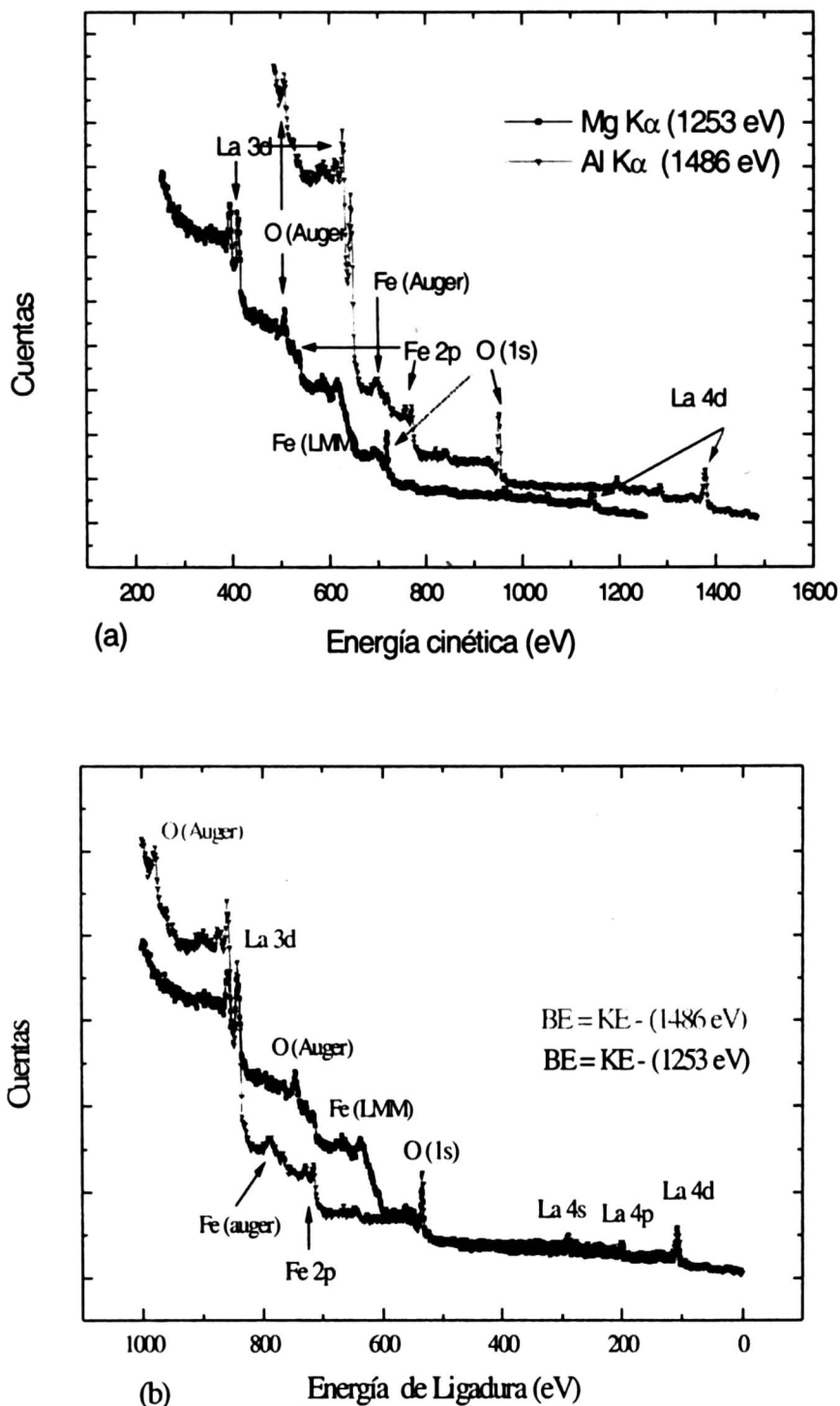


Figura 4. Espectros XPS de una muestra de LaFeO_3 donde se presenta el número de cuentas o intensidad en función de (a) la energía cinética y en función de (b) la energía de ligadura BE. En (a) coinciden los picos Auger y en (b) los picos de fotoelectrones.

Espectroscopía de Electrones Auger (AES)

En esta técnica la muestra se irradia con electrones monocromáticos y se cuentan los electrones Auger producidos o emitidos por la muestra. El resultado directo también es el número de cuentas $N(K)$ en función de la energía cinética K de los electrones. Pero para determinar la posición los picos correspondientes a electrones Auger se presenta la derivada del número de cuentas, dN/dE . En los espectros Auger sólo se tienen picos Auger y no de fotoelectrones [2].

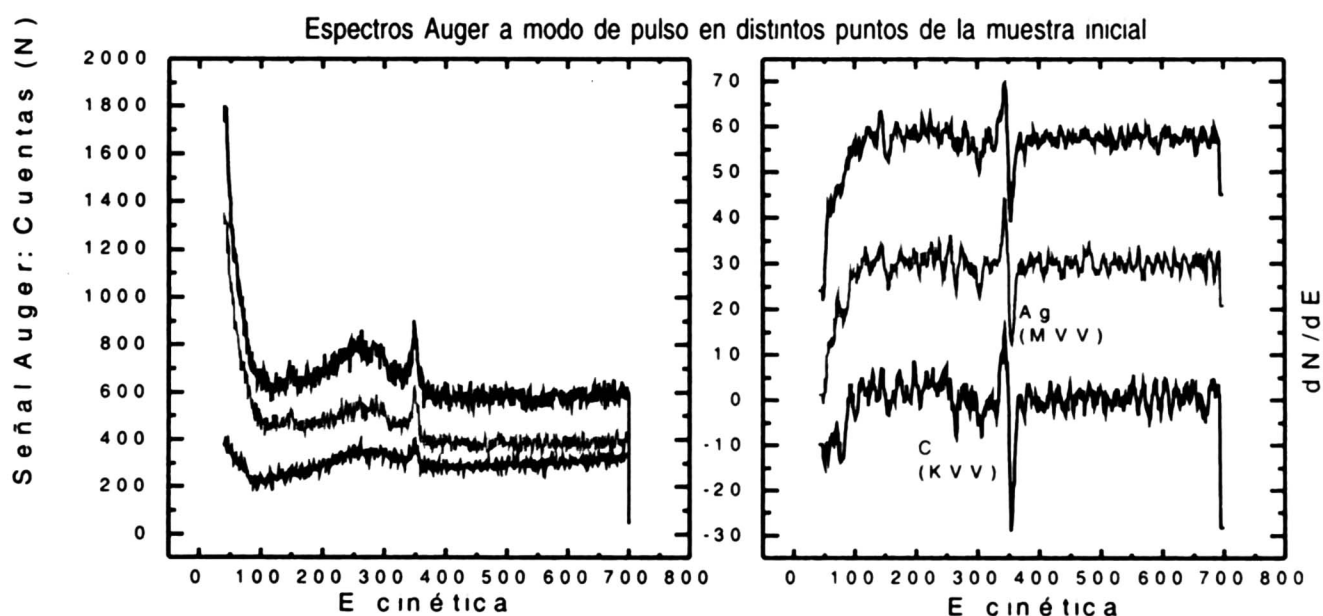


Figura 5. Espectros AES en modo de pulso de una muestra a analizar en condiciones iniciales y como se ve está contaminada con Ag y C.

Óxidos de los metales de transición 3d

Los espectros de transiciones Auger para bajas energías de la mayoría de los óxidos de elementos de bajo número atómico muestran líneas Auger en energías más bajas que la línea Auger de los elementos puros.

Los espectros de transiciones Auger para bajas energías de compuestos de metales de transición generalmente muestran dos transiciones Auger de intensidades comparables (Figura 6), una de ellas aparece a una energía más alta que la línea Auger del metal puro.

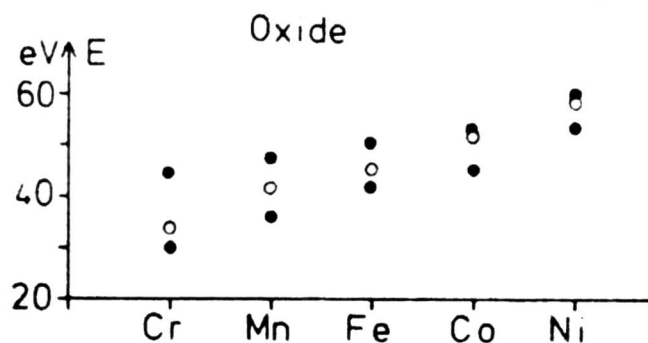


Figura 6. Energías Auger de óxidos de metal de transición 3d. Círculos vacíos: elementos puros, círculos llenos: compuestos.

Este comportamiento puede ser explicado por un simple modelo de transición cruzada. Los niveles de valencia de los átomos de metales de transición consisten de capas 3d parcialmente llenas y capas 4s completamente llenas, las que en un sólido forman la banda de valencia. Aplicando el modelo de banda iónica (Figura 7a) los niveles de oxígeno 2p se llenan con dos electrones 4s del metal.

Las dos líneas Auger observadas son 1) transiciones Auger de los niveles 3d afectados por un corrimiento químico y 2) transiciones cruzadas desde los niveles 2p del oxígeno a huecos en las capas internas del metal. Una comparación con las bandas de valencia de los óxidos muestra que el modelo puede ser aplicado: en el estado oxidado la banda 4s-3d de un metal se separa en una banda vacía 4s y una banda de valencia 3d parcialmente llena al nivel de Fermi (Figura 7b). Adicionalmente se observa otra banda debido al nivel 2p del oxígeno. Con éste esquema de energía las transiciones Auger pueden explicarse de la misma manera que en el modelo iónico. La separación de energía de los dos picos Auger, uno por corrimiento químico y el otro por transición cruzada, corresponden a dos veces la banda prohibida. En el caso de los óxidos la línea de separación observada decrece con el incremento del número atómico de acuerdo con cálculos de estructuras de bandas.

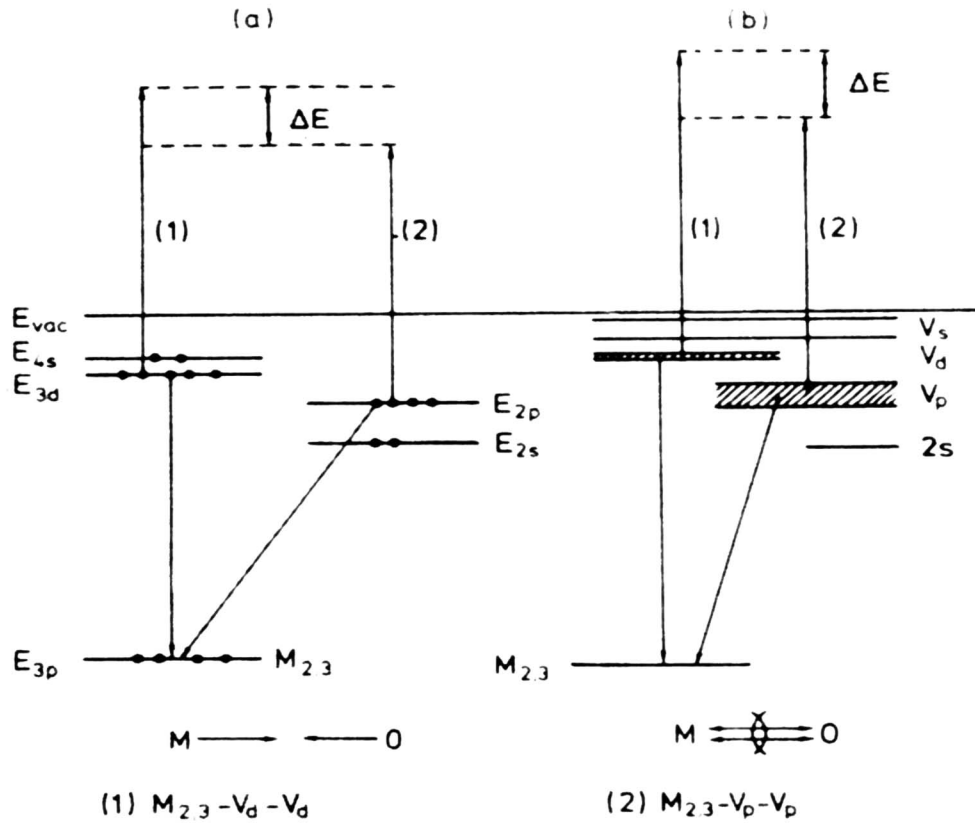


Figura 7. Esquema de los niveles de energía de las transiciones Auger de óxidos de metales 3d: (a) modelo iónico; (b) modelo de banda.

Sputtering y Profundidad (Depth Profiling)

En algunas técnicas en estudio de superficies se hace incidir sobre la muestra iones de gases inertes (He^+ , Ne^+ o Ar^+) en un rango de energía desde unos pocos eV hasta algunos keV. Si el ión incidente de energía E_0 y masa M_1 es dispersado por un átomo de masa M_2 de la superficie bajo un ángulo θ_1 (Figura 8), entonces el ión dispersado tiene una energía E_1 dada por

$$\frac{E_1}{E_0} = \frac{1}{(1+A)^2} [\cos\theta_1 \pm (A^2 - \sin^2\theta_1)^{1/2}]^2 \quad (3)$$

donde $A = M_2/M_1$ y el signo positivo es para $A > 1$, ambos signos para $A < 1$. Similarmente el átomo dispersor gana una energía E_2 . Asumiendo que está inicialmente en reposo y es arrancado a un ángulo θ_2 relativo a la trayectoria del ión incidente, E_2 está dado por:

$$\frac{E_2}{E_0} = \frac{4A}{(1+A)^2} \cos 2\theta_2 \quad (4)$$

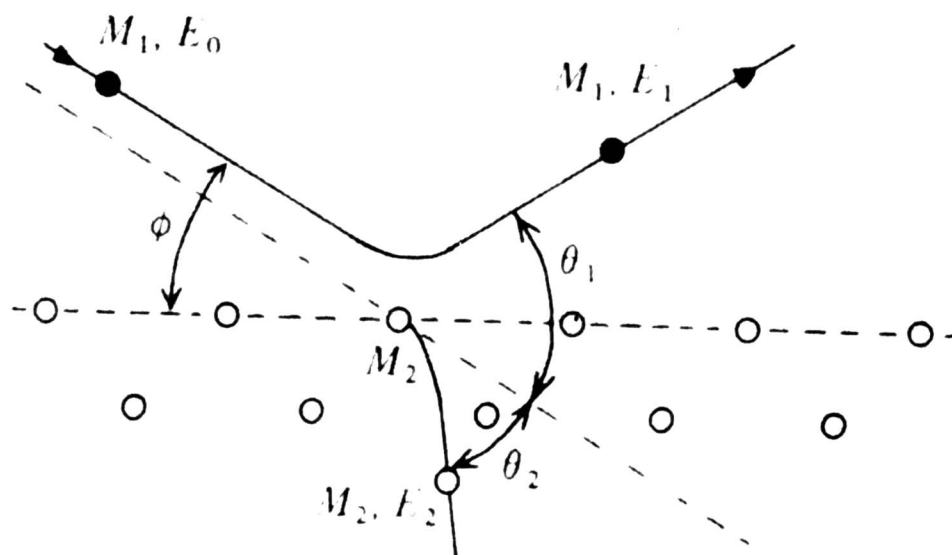


Figura 8. Representación esquemática de la dispersión de unión por átomos de la superficie.

La energía E_2 puede ser una fracción importante de la energía del ión incidente, excediendo ampliamente la energía de ligadura local del átomo dentro del sólido, y entonces el átomo puede colisionar con muchos otros produciendo una cascada de colisiones en las que una cantidad importante de átomos del sólido son puestos en movimiento. Dependiendo de la energía cinética adquirida, algunos de estos átomos pueden ser arrancados de la muestra produciendo la erosión de la superficie. Este fenómeno es conocido como «sputtering» y se utiliza ampliamente para limpiar la superficie de contaminantes. Típicamente, un estudio de la composición química en función de la profundidad (depth profiling) consiste en erosionar la superficie mediante el bombardeo con iones de Ar^+ a una energía de 2 KeV e ir midiendo espectros XPS o AES.

Multicapas magnéticas de manganita/titanato de estroncio

En este trabajo se analizaron dos muestras de multicapas constituidas por una película de manganita otra de titanato de estroncio y una más de la misma manganita con diversos espesores sobre sustratos de MgO en un caso y SrTiO_3 en otro. La manganita en cuestión es $\text{La}_{0.55}\text{Sr}_{0.45}\text{MnO}_3$ la cual tiene una estructura tipo perovskita (Figura 9), con distorsión romboédrica por el dopaje de Sr ($x = 0.45$) [3]. Los átomos de Mn están en el centro de un octaedro y los de oxígeno en los vértices del mismo.

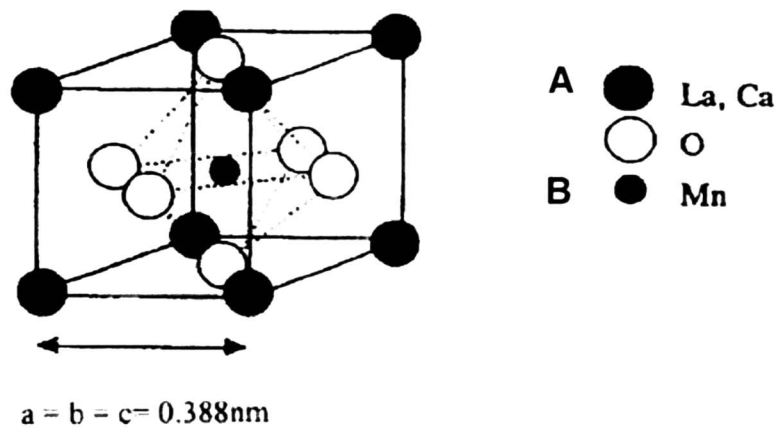


Figura 9. Estructura perovskita cúbica ideal ABO_3 . A es un catión grande (La, Ca), similar en tamaño al O^{2-} ; B es el catión pequeño Mn^{3+} o Mn^{4+} en coordinación octaédrica con el oxígeno.

La presencia de átomos magnéticos (Mn) rodeados por átomos no magnéticos (O) en las manganitas $La_{1-x}Sr_xMnO_3$, da lugar a dos tipos predominantes de intercambio que compiten entre sí: el *superintercambio*, que favorece un orden antiferromagnético (AF) entre iones Mn, y el *doble intercambio*, que favorece un orden ferromagnético (FM). Los oxígenos están siempre en el estado O^{2-} , por lo que el estado de ionización de los Mn se determina por la proporción x de dopante. En el extremo $x = 0$, se tienen todos Mn^{3+} y al aumentar x , aparece la misma concentración x de Mn^{4+} [4].

Las manganitas presentan diagramas de fase complejos, ya que sufren transiciones de fase magnéticas, estructurales y en sus propiedades de transporte eléctrico. En la figura 10 se muestra el diagrama de fases para monocristales de $La_{1-x}Sr_xMnO_3$ al variar la temperatura y la concentración de dopante x .

Para x cercana a cero, el sistema es siempre aislante y presenta un orden antiferromagnético cantedo a temperaturas bajas. Para x un poco mayor se observa un orden ferromagnético a bajas temperaturas, y el sistema sigue siendo aislante en todo el rango de temperaturas. Para $0.15 < x < 0.5$, se observa una transición metal-aislante en todo el rango de temperaturas y para temperaturas altas además hay una transición paramagneto-ferromagneto [5].

En nuestro caso con $x = 0.45$ tenemos la manganita en fase ferromagnética-metálica.

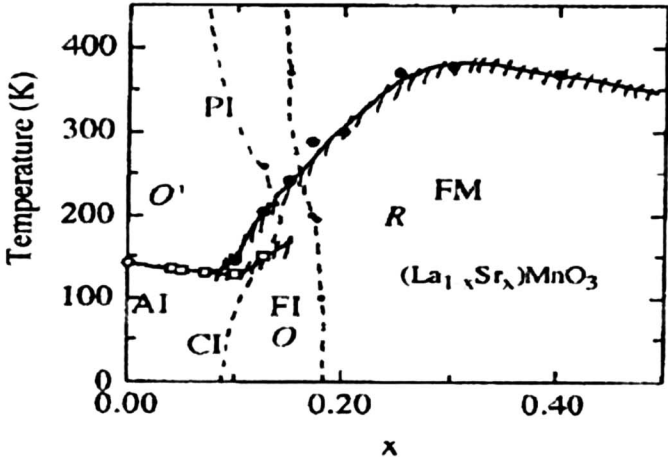


Figura 10. Diagrama de fases para monocristales de $\text{La}_{1-x}\text{Sr}_x\text{MnO}_3$. O' y O representan una estructura ortorrómbica, y R una romboédrica. Se observan distintas fases magnéticas: A (antiferromagnética), C (canteda), F (ferromagnética) y P (paramagnética). Hay regiones donde el material es metálico (M) y otras donde es aislante (I).

Crecimiento de las tricapas

Para éste trabajo las películas delgadas fueron crecidas por sputtering DC a partir de un blanco cerámico de $\text{La}_{0.55}\text{Sr}_{0.45}\text{MnO}_3$ para la capa de la manganita y otro blanco de SrTiO_3 para la capa central. Dentro de la cámara de vacío estaban los dos blancos. Primero se trabaja el sustrato con el blanco de manganita por un tiempo determinado considerando que la tasa de deposición es $1.5 \Rightarrow/\text{min}$, luego se traslada el sustrato con su primera película ante el blanco del titanato para depositar la capa central y luego otra vez ante el cerámico de manganita para depositar la última capa. Durante la deposición el sustrato se mantiene a una temperatura de 700°C .

Los sustratos utilizados son monocristales de SrTiO_3 (100) que tiene también una estructura tipo perovskita con parámetro de red $3.905 \Rightarrow$, 0.9% diferente al de la manganita ($3.869 \Rightarrow$) y MgO (100) el cual tiene una estructura tipo NaCl con un parámetro de red de $4.216 \Rightarrow$. Sobre el sustrato de SrTiO_3 el crecimiento de la manganita sería celda sobre celda por los parámetros de red similares, distorsionando levemente el tamaño de la celda para subsanar la diferencia entre los parámetros de red (*lattice mismatch*) estirándose el film de manganita en el plano del sustrato y contrayéndose en la dirección perpendicular. Sobre el sustrato de MgO la interfase será mucho más desordenada ya que los parámetros de red

difieren en casi un 10% y el sistema deberá introducir un gran número de dislocaciones para relajar las tensiones.

Los espesores nominales de las películas están dados en la siguiente tabla:

Tabla 1. Características de las muestras.

Muestra	substrato	Primera capa / espesor (\Rightarrow)	Segunda capa / espesor (\Rightarrow)	Tercera capa / espesor (\Rightarrow)
Muestra 1	MgO	$\text{La}_{0.55}\text{Sr}_{0.45}\text{MnO}_3$ / 400	SrTiO_3 / 20	$\text{La}_{0.55}\text{Sr}_{0.45}\text{MnO}_3$ / 500
Muestra 2	SrTiO_3	$\text{La}_{0.55}\text{Sr}_{0.45}\text{MnO}_3$ / 400	SrTiO_3 / 50	$\text{La}_{0.55}\text{Sr}_{0.45}\text{MnO}_3$ / 500

La muestra 1 además fue terminada con litografía óptica y grabado iónico para producir los contactos eléctricos de un diseño donde se pretendía medir la conducción en la juntura a través de la capa central de SrTiO_3 que es aislante y de las capas conductoras de manganita. Entonces la tricapa no cubre todo la superficie de la muestra, sino sólo dos pequeñas regiones de forma casi triangular (Figura. 11).

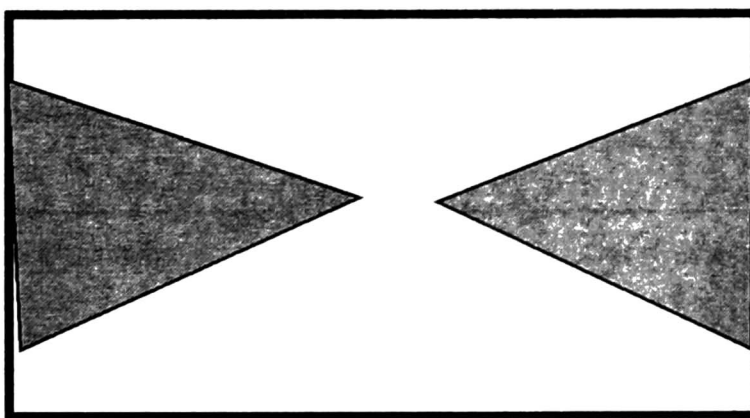


Figura 11. Esquema de la superficie de la muestra 1. Las dos regiones triangulares sombreadas son las tricapas (tabla 1). El substrato es MgO.

Resultados y Discusión

Las técnicas empleadas para analizar químicamente la superficie de la muestra fueron: Espectroscopía de Fotoemisión de rayos X (XPS) que analiza toda la superficie de la muestra (unos 20 mm²) y Espectroscopía de Electrones Auger (AES) que analiza sólo unos 0.5mm² de la superficie de la muestra.

En los primeros espectros Auger en modo pulso para no cargar eléctricamente la muestra, se aprecian picos de C y Ag poco claros (Figura 5) lo que indica que sobre la superficie se había depositado Ag y estaba contaminada con C. Entonces se había depositado Ag sobre la muestra para una caracterización (microscopía) anterior al estudio con espectroscopia de electrones.

Luego se «bombardeó» la muestra con argón en varias etapas, controlando parámetros como corriente (con una copa de Faraday), presión y tiempo para conocer la tasa de «grabado iónico». La corriente inicial para el bombardeo de Ar^+ fue de 3.0×10^{-8} A, correspondiente a 3.8 mA/cm^2 , a una presión de 1.0×10^{-6} torr, y un potencial de 2.6 kV. Luego de 135 segundos en el espectro Auger se observan sólo picos de Ag. Seguimos bombardeando la muestra con las mismas condiciones en intervalos de tiempo de 5 ó 10 minutos, observando en los espectros que el pico correspondiente a Ag va decreciendo para desaparecer totalmente después de unos 60 minutos, a la vez que aparecen los picos de Mn, La y O en espectros AES, los que se hicieron en distintos puntos de la muestra (Figura 12), observándose que hay presencia de LaSrMnO_3 fuera de las regiones triangulares. Esto indica que el litografiado óptico resultó defectuoso ya que sólo debería haber manganita en la regiones triangulares.

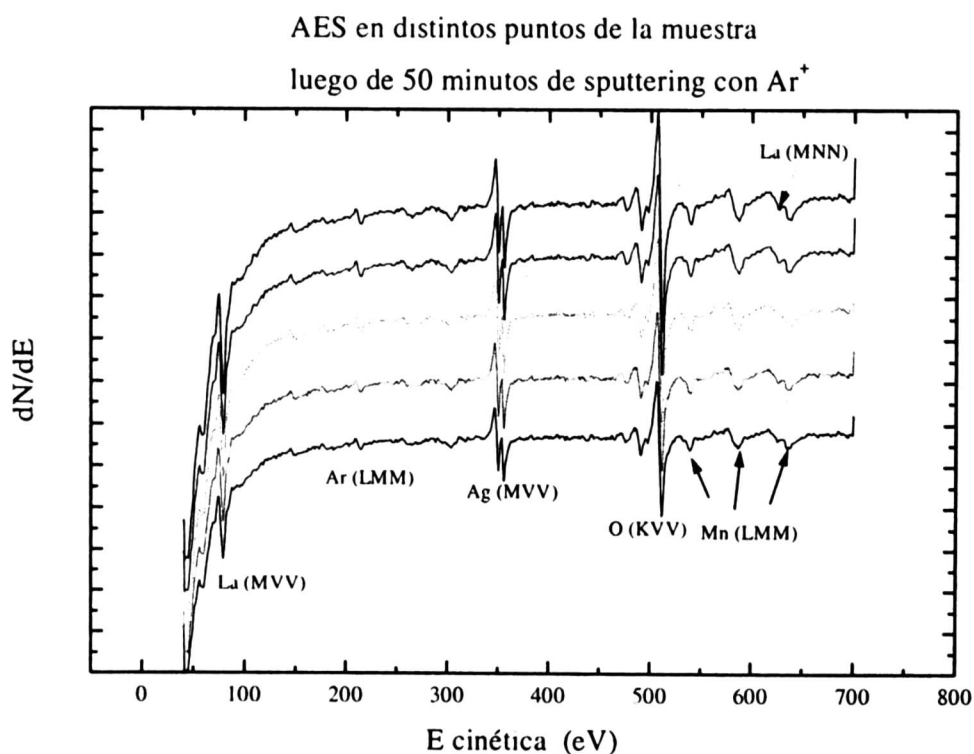


Figura 12. Espectros AES de la muestra 1 luego de ser «bombardeada» por 50 minutos. Se observa que hay manganita en toda la superficie, es decir, también fuera de las regiones triangulares.

También se obtuvieron luego de 50 minutos de bombardeo iónico espectros XPS (Figura 13). En ellos se observa además el pico de fotoemisión 3d del Sr. En el espectro AES no se apreciaba la línea Auger asociada al Sr por el rango de energía de la medición.

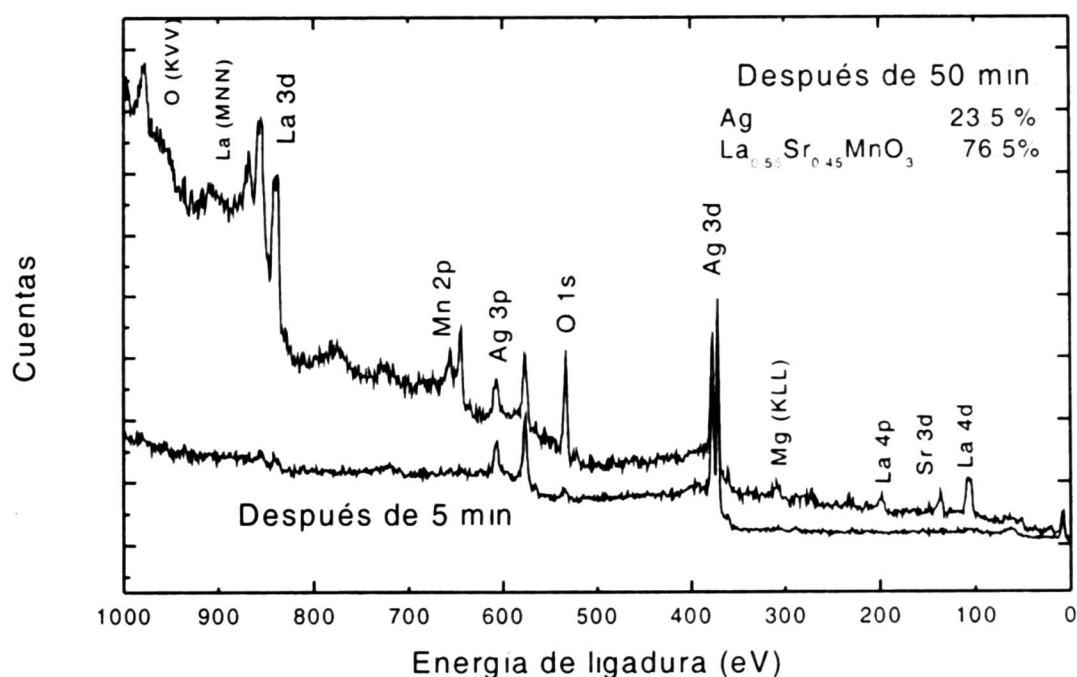


Figura 13. Espectros XPS, con ánodo de AlK α (1486.6 eV), de la muestra 1 a los 5 y 50 minutos luego de ser bombardeada con Ar $^{+}$.

Se siguió bombardeando y tomando espectros periódicamente. Luego de más de 3 horas de dicho proceso ya no se observa manganita en los espectros AES fuera de las regiones triangulares y se aprecia el pico Auger del oxígeno indicando que el sustrato no sería SrTiO $_3$ sino posiblemente MgO por lo cual se tomó un espectro XPS con un rango de energía mayor (energía de ligadura desde 1400 a 0 eV) y con el ánodo de Al para usar la línea Al K α como fuente de fotones y así poder observar si hay pico de Mg. Efectivamente obtuvimos el resultado esperado, se encontró el pico de Mg. (Figura 14)

Con espectroscopia Auger se analizaron diferentes zonas de la muestra encontrándose que el sustrato era MgO y verificando que las tricapas eran sólo dos regiones casi triangulares de toda la superficie de la muestra, es decir, los espectros Auger en zonas de las regiones triangulares acusan presencia de LaMnO y fuera de ellas de MgO. Se midieron entonces las intensidades de los picos de los elementos presentes

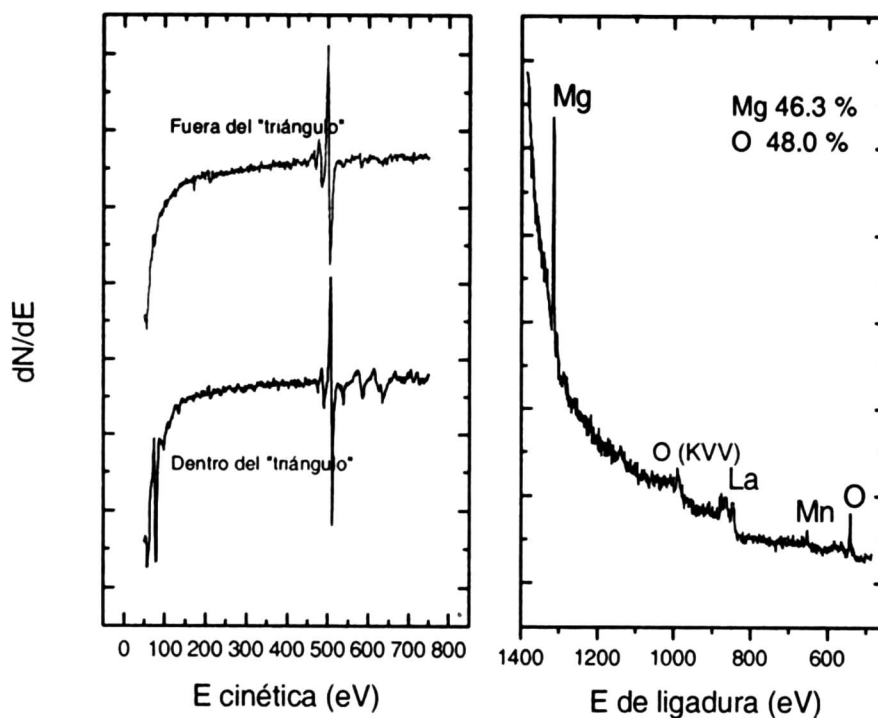


Figura 14. Espectros AES a la izquierda y XPS en la derecha de la muestra 1 luego de 194 minutos de bombardeo. Se observa que fuera de la región triangular ya no hay manganita y de XPS con un rango ampliado de energía se tiene un pico de Mg correspondiente al sustrato MgO.

en los espectros AES tomados en posiciones conocidas a lo largo de una línea recta, dirección z, en la superficie de la muestra (Figura 15). Se encuentra un buen acuerdo con la forma de la muestra, es decir, dentro de la región triangular aparecen los picos de La y Mn y fuera de ella no se presentan.

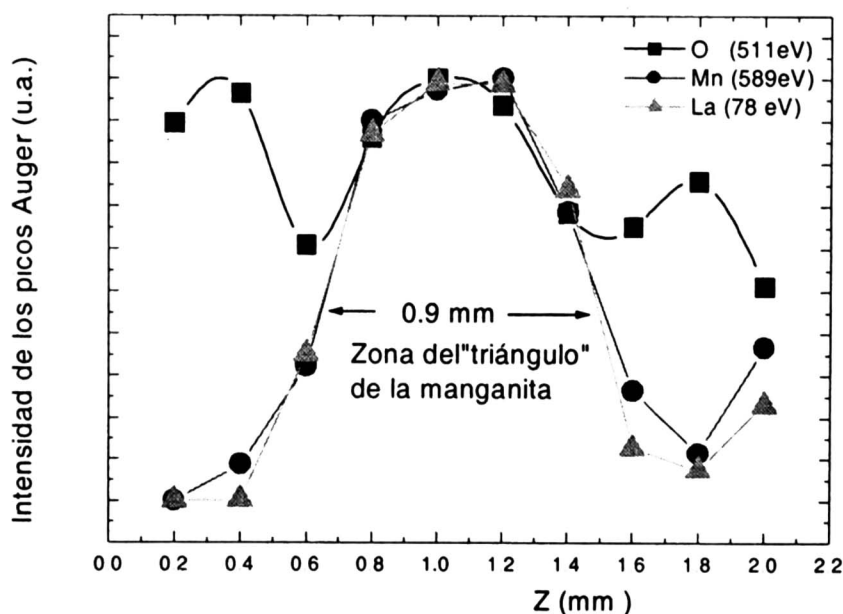


Figura 15. Intensidad relativa de los picos Auger de algunos elementos en función de la dirección Z sobre la superficie de la muestra 1.

Cambiando el potencial de bombardeo de 2.6 a 3 kV, se dieron tiempos de bombardeo mayores y se tomaron espectros AES de una región triangular, observándole siempre la manganita. Después de un total de 760 minutos de bombardeo iónico se observó en los espectros AES que los picos de La y Mn no estaban presentes sino sólo el oxígeno, entonces se tomó un espectro XPS (Figura 16) y resultó que sólo se tiene MgO, es decir habíamos llegado al substrato sin encontrar la capa de SrTiO₃.

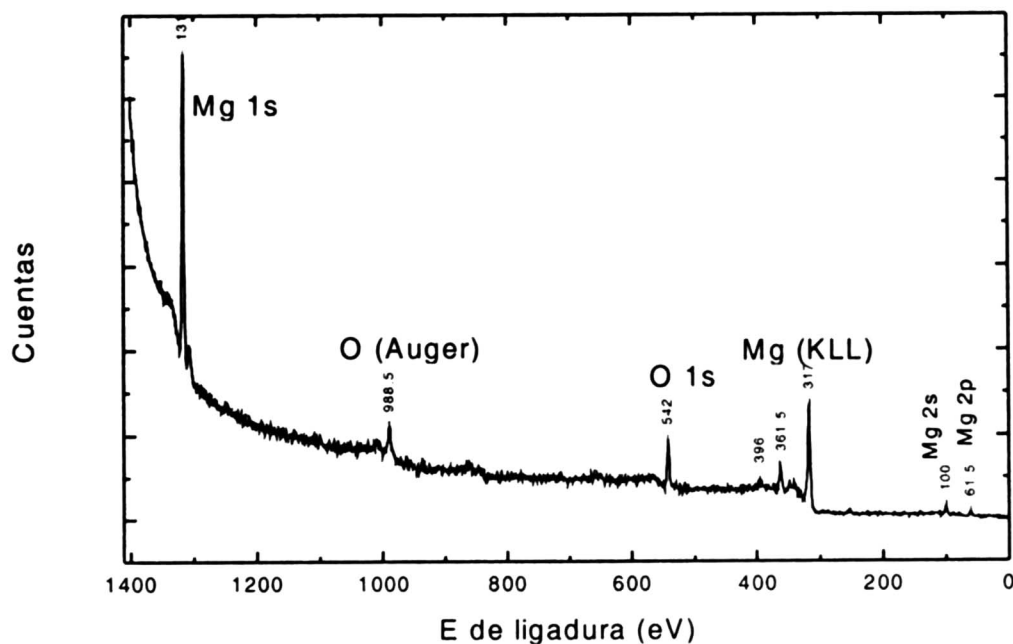


Figura 16. XPS de la muestra 1 al final de 671 minutos de bombardeo. Se llegó al substrato MgO sin lograr ver la capa de SrTiO₃.

Conociendo el espesor nominal de la tricapa, 920 \Rightarrow , y del tiempo que se tardó en eliminar la capa de Ag que tenía encima, 60 min, se estimó una tasa media de grabado iónico de 1.5 \Rightarrow /min.

El espectro XPS realizado a una muestra de SrTiO₃ se presenta en la Figura 17 y es lo que esperaríamos tener de la capa central.

Para la muestra 2, se tomaron también espectros iniciales de XPS y AES en los cuales se encontró un pico de C. Luego de 5 minutos de bombardeo a una corriente de 3.0×10^{-8} A, correspondiente a 3.8 mA/cm² y una presión para el Ar⁺ de 1.0×10^{-6} torr, y un potencial de 2.6 kV se había logrado eliminar el pico de C (Figura 18), quedando sólo los correspondientes a la manganita.

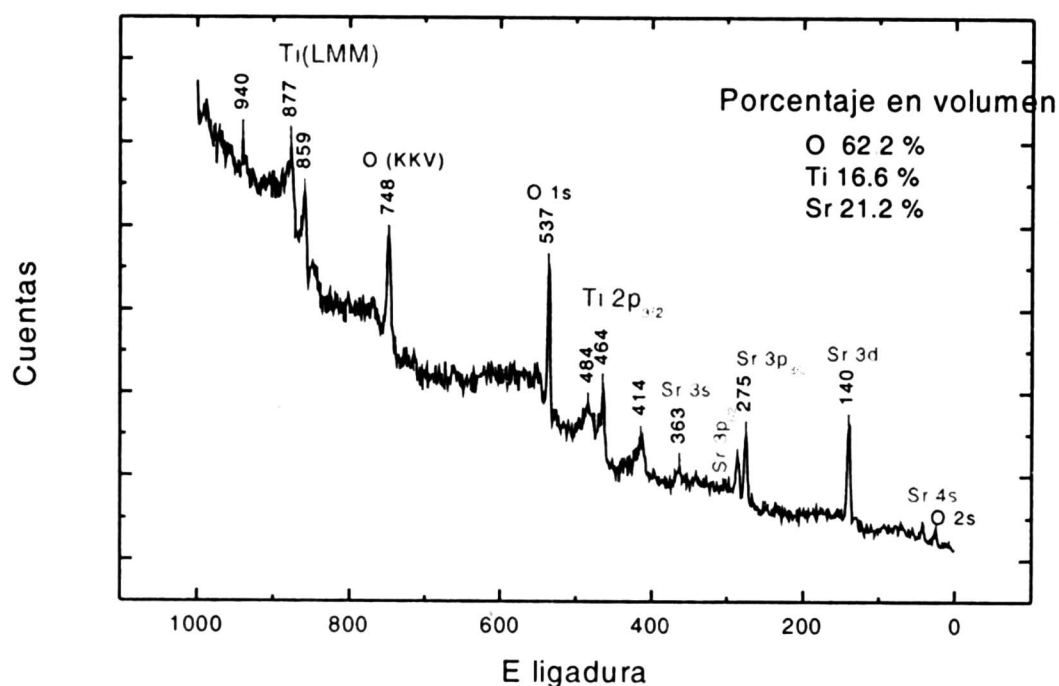


Figura 17.- XPS de una muestra patrón de SrTiO_3 irradiada con ánodo de $\text{MgK}\alpha$.

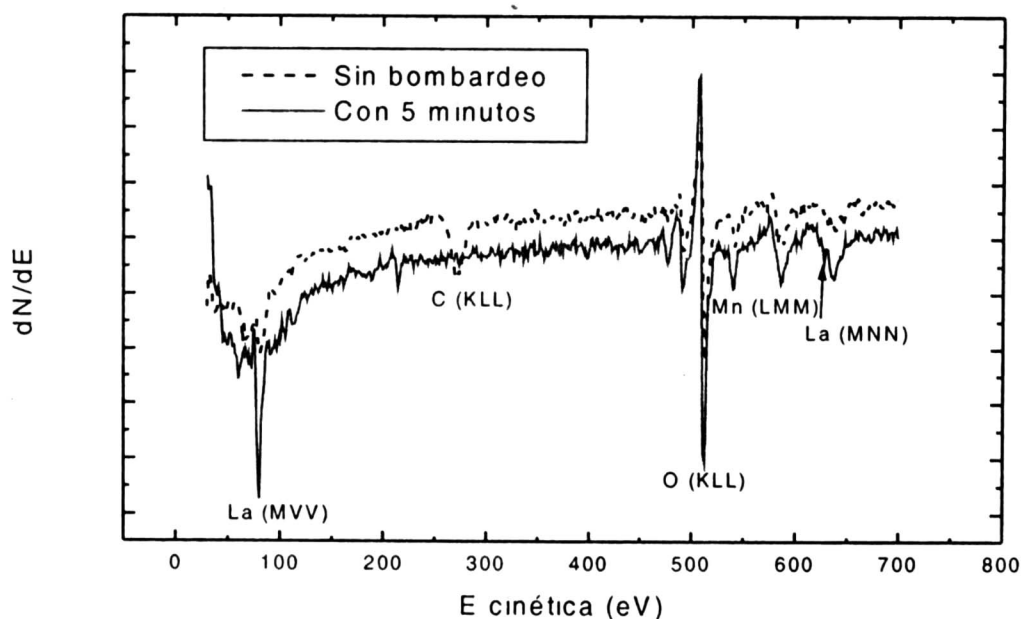


Figura 18. AES de la muestra 2 al inicio y luego de 5 minutos de bombardeo. Se observa que la muestra estaba contaminada con C y luego sólo se tiene la manganita.

Se bajó entonces el potencial a 2.1 kV y se dieron tiempos de bombardeo inicialmente de 60 minutos hasta 4 horas y luego cada 10 minutos. Después de un total de 615 minutos y dado que los espectros XPS no cambiaban, se aumentó el potencial para el bombardeo a 3 kV. Luego de 130 minutos en estas condiciones se empezó a observar un

pequeño pico de Ti. Se bombardeó entonces cada 5 minutos y se realizó espectros AES a la muestra 2. El pico de Ti apenas aumentó y los picos de la manganita casi no variaron para finalmente quedar los de la manganita y desaparecer los de Ti (Figura 19). No se logró tener un espectro sólo de SrTiO_3 como en la figura 16, sino que siempre se observan los picos de manganita.

En la Figura 19 se muestra las intensidades relativas de los picos tomados de los espectros AES para la zona que correspondería a la capa central. Aquí se puede apreciar un leve crecimiento del pico de Ti que luego vuelve a desaparecer. Los picos de Mn y La apenas variaron. Los tiempos están dados a partir del momento en que se apreció el primer pico de Ti.

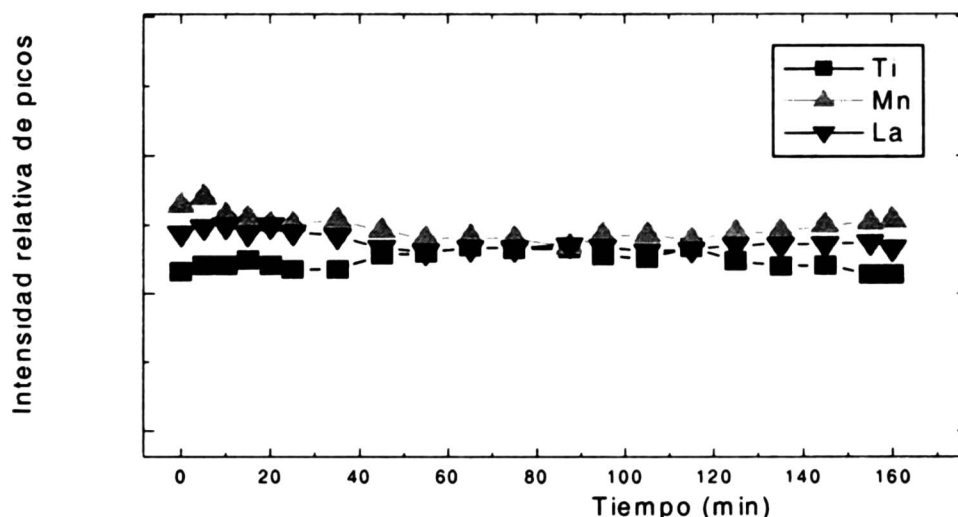


Figura 19. Intensidad relativa de los picos tomados para dos elementos de la manganita (Mn y La) y del Ti en función de los tiempos de bombardeo a partir del momento en que empieza a apreciarse el pico de Ti.

Conclusiones

- De la muestra 1 y la gráfica de la figura 15 se obtuvo el ancho del haz de electrones, con que se irradia la muestra para obtener espectros AES en el equipo utilizado, resultando éste de 0.5 mm. También se encontró que el litografiado óptico para esta muestra había sido defectuoso porque encontramos manganita fuera de las regiones triangulares.

- No se pudo observar los picos correspondientes a la capa central (SrTiO_3) de las tricapas. Como en las tres capas hay Sr y O entonces interesaba observar los picos de Ti, los que no pudieron apreciarse debido al espesor de la capa central ($20 \Rightarrow$ en la primera muestra y $50 \Rightarrow$ en la segunda muestra). Esto se corrobora también simulando el bombardeo iónico con los parámetros usados en el análisis de las tricapas. El resultado de dicha simulación se presenta en la Figura 20. Como se ve, la muestra es dañada significativamente para espesores menores a $100 \Rightarrow$.
- Para poder analizar la composición química dentro de la capa central para espesores menores a $100 \Rightarrow$ se sugiere hacer sólo una bicapa y analizarla con XPS/Auger (sin grabado iónico) inmediatamente después de haberla fabricado.

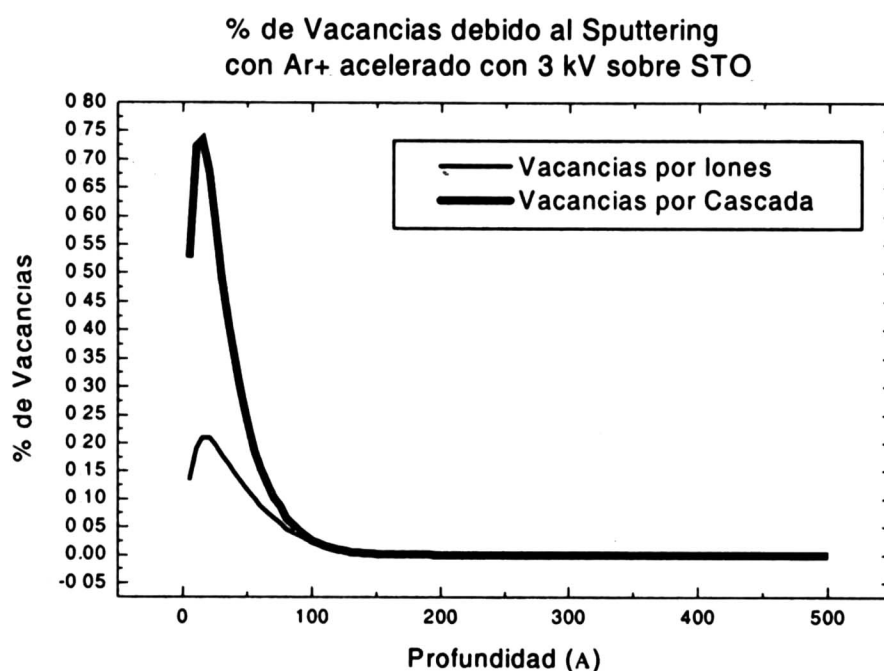


Figura 20. Simulación en el trem del porcentaje de vacancias en el SrTiO_3 por bombardeo de Ar^+ acelerados con 3kV.

Referencias

- [1] *Modern Techniques of Surface Science*, D.P. Woodruff & T.A. Delchar. Cambridge Solid State Science Series
- [2] R. Weiismann y K. Muller, *Auger Electron Spectroscopy – A Local Probe for Solid Surfaces*. Surface Science Reports 105 (1981)
- [3] Coey et al., *Adv. Phys.* 48 vol.2 (1999)

- [4] *M. Granada, Transporte Eléctrico en Películas Delgadas de $\text{La}_{0.6}\text{Sr}_{0.4}\text{MnO}_3$: Magnetoresistencia y Efecto Hall, Tesis de Licenciatura en Física, Instituto Balseiro (2001)*
- [5] *Colossal Magnetoresistive Oxides, Y. Tokura. Advances in Condensed Matter Science.*

Agradecimientos

Quiero agradecer en primer lugar al Instituto Balseiro por haberme dado la oportunidad de trabajar en sus laboratorios y compartir con sus investigadores durante este mes de febrero. Una experiencia enriquecedora tanto para mi desarrollo académico como personal. Gracias a Hugo Ascolani por todo el interés mostrado en enseñarme lo que sé sobre física de superficies y por permitirme y ayudarme a manejar el equipo de XPS y AES. A Laura Steren, quién nos proporcionó las muestras y el trabajo, como aquella vez que trabajamos juntos en noviembre del 2002 cuando asistí a la Escuela de Física del Sólido. A Esteban por sus discusiones sobre sputtering y superficies. A los técnicos de la división de colisiones atómicas y a mis compañeros becarios por todo su apoyo. Por supuesto a nuestra Facultad de Ciencias de la UNI, a mi asesor A. Gutarra y al profesor W. Estrada por sus cartas de recomendación.

Revistas publicadas

VOLUMEN	NÚMERO	Mes de Publicación	AÑO
1	1	Junio	1995
2	1	Julio	1996
2	2	Noviembre	1996
2	3	Diciembre	1996
3	1	Setiembre	1998
4	1	Octubre	1999
5	1	Julio	2001
6	1	Enero	2002
6	2	Abril	2002
7	1	Febrero	2003
7	2	Setiembre	2003

Eventos

1. VI Seminario Internacional de Optimización y Áreas Afines. Lima, 6-10 de octubre del 2003. E-mail: imca@uni.edu.pe, sosa@uni.edu.pe
2. Simposio Nacional de Química e Ingeniería Química. Lima, 17-18 de octubre del 2003. Web: www.sqp.8k.com
3. X Simposio Peruano de Energia Solar (X SPES) y Seminario Internacional sobre Tecnologías Económicas para la Descontaminacion y Desinfección de Agua. Cuzco, 17-22 de noviembre del 2003. Web: www.unsaac.edu.pe/eventos/xspes
4. III Curso Taller Sobre «Adsorbentes para la Protección Ambiental». Lima, diciembre del 2003. E-mail: agutarra@uni.edu.pe
5. Sexto Encuentro de Física. Lima, 5-7 de enero del 2004. Web: www.encuentrodefisica.org
6. Quinto Simposio Latinoamericano de Física de Altas Energías (V SILAF AE). Lima, 12-16 de julio del 2004. Web: www.v-silafae.org

Fé de Erratas

Página	Renglón	Dice	Debe decir
4	7	exelente	excelente
7	1	unitaria	unitario
7	5	Pincaré	Poincaré
8	15	planta	plana
14	9	relacions	relaciones
17	13	asocida	asociada
18	5	elípticas	hiperbólicas
18	9	demuestra	demuestra
82	9	este	esta
82	13, 14	\Rightarrow	Å
83	6	\Rightarrow	Å
92	6,12,13,14	\Rightarrow	Å
93	Tabla 1	\Rightarrow	Å

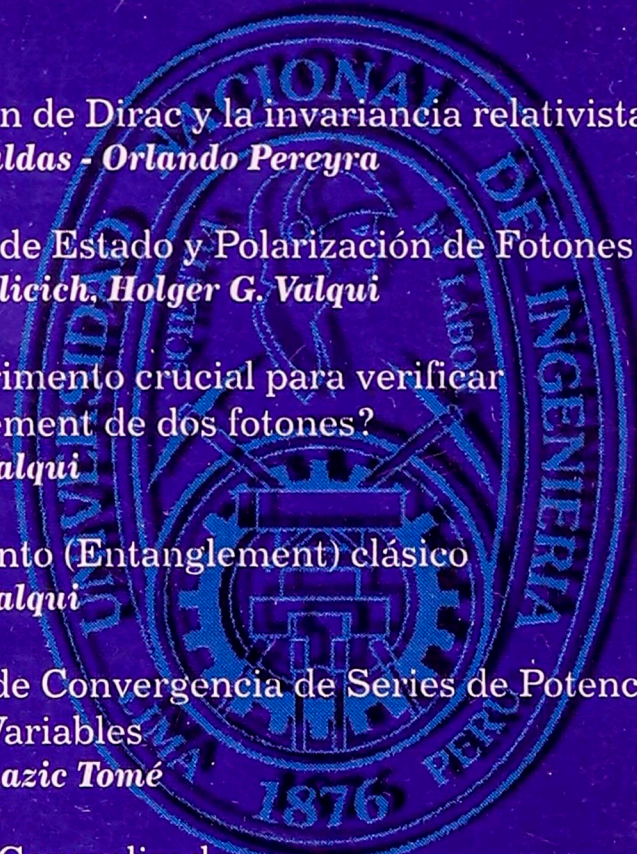
 edición, diagramación e impresión
ditorial HOZLO S.R.L.

Psje. Santa Rosa 191-501
Lima - Perú
Telefax: 428-4071
e-mail: GUZLOP@terra.com.pe

Esta edición consta de 0500 ejemplares

Lima, septiembre del 2003

ÍNDICE

- 
- La ecuación de Dirac y la invariancia relativista
Wilfredo Caldas - Orlando Pereyra 5
 - Funciones de Estado y Polarización de Fotones
Oscar Miculicich, Holger G. Valqui 19
 - ¿Un Experimento crucial para verificar
el entanglement de dos fotones?
Holger G. Valqui 28
 - Enlazamiento (Entanglement) clásico
Holger G. Valqui 33
 - Dominios de Convergencia de Series de Potencias
de Varias Variables
Renato Benazic Tomé 45
 - Gradiente Generalizado
Joseph Wilmer Simpe Laura, Mg. William Echegaray Castillo 59
 - Estudio de la Composición Química de Multicapas
Magnéticas en Función de la Profundidad
J. Carlos Rojas Sánchez - Hugo Ascolani - Laura Steren 81