

# Revista de la Facultad de **CIENCIAS** de la UNI **REVCUNI**

UNIVERSIDAD NACIONAL DE INGENIERÍA

LIMA - PERÚ

Editor: *Armando Bernui Leo*

Comité editorial:

*Félix Escalante Del Águila*  
*Manfred Horn*  
*Holger Valqui Casas*

Oficina de publicación:

*Facultad de Ciencias - UNI*  
Av. Túpac Amaru 210 - Rímac  
Teléfono: (0051) (1) 481-0824  
E-mail: [bernui@uni.edu.pe](mailto:bernui@uni.edu.pe)

Ley del Depósito Legal N° 15010199-3597



Revista de la Facultad de  
**CIENCIAS** de la UNI  
**REVCUNI**

# Presentación

---

*Me complace hacerles llegar un número más de nuestra REVCIUNI. Se trata del quinto número desde que me encargaron la labor de editor a fines de 1999.*

*La oportunidad es propicia para un par de reflexiones. Han sido años difíciles porque en nuestra facultad no abundan los autores de artículos científicos. Parte de nuestras energías se fueron tratando de convencer a nuestros colegas de la necesidad de mantener el ritmo de una publicación periódica semestral para nuestra REVCIUNI. Ustedes podrán comprobar que en estos tres años la mayoría de artículos han sido escritos por nuestros jóvenes pre-docentes, bachilleres e incluso algunos de nuestros estudiantes. Lo que ocurre es que redactar un artículo científico es una tarea difícilísima porque demanda dedicación y paciencia además de conocimiento del tema.*

*Sin embargo, nuestro esfuerzo de alguna manera ha servido para que nuestra revista se convierta en un medio de transmisión de información científica (a nivel divulgativo y a nivel especializado) entre los miembros de nuestra facultad.*

*Espero que mi sucesor encuentre un camino más o menos allanado, ya que ahora son más los miembros de nuestra facultad que reclaman – sino como autores al menos como lectores– la aparición del próximo número de nuestra revista.*

*Quedo muy agradecido a todos aquellos que me brindaron su apoyo.*

**Armando Bernui**  
El editor.

# Raíces y álgebra de Lie semisimples diagramas de Dynkin

---

*Aldo Arroyo Montero\**

## *RESUMEN*

En el presente trabajo desarrollamos la teoría a las álgebras de Lie. Empezamos dando el concepto de álgebra de Lie, luego nos centramos en el análisis de las denominadas álgebras de Lie semisimples, dentro de este contexto definimos el concepto de raíz simple. Al final mostramos dos ejemplos de álgebras de Lie semisimples, el álgebra  $su(2)$  y el álgebra  $su(3)$ , las cuales son muy usadas en la Física, por ejemplo el álgebra  $su(2)$  viene a ser el álgebra del operador del momento angular.

## *ABSTRACT*

We study here the theory of Lie algebras. We start giving the concept of the Lie algebras, then we analyze the semisimple Lie algebras, then we also define the concept of simple root. At the end we show two examples of semisimple Lie algebras, the algebra  $su(2)$  and the algebra  $su(3)$ . Those algebras are very used in physics, for example the algebra  $su(2)$  is the algebra of the angular momentum operator.

---

\* Grupo de Física Teórica, Fac. de Ciencias, UNI.  
J19970390@uni.edu.pe

## Introducción

El primer ejemplo comúnmente usado en el estudio de las álgebras de Lie viene a ser el álgebra  $su(2)$ , la base usual para esta álgebra viene a ser el conjunto de vectores  $\{S_1, S_2, S_3\}$  donde las relaciones de conmutación vienen dadas por

$$[S_i, S_j] = \sum_{k=1}^3 i\epsilon_{ijk} S_k$$

También existe otra base de vectores para esta álgebra  $\{S_{\pm}, S_3\}$ , denominada base esférica [1] o cíclica. La relación que existe entre ambas bases viene dada por

$$S_3 = S_3$$

$$S_{\pm} = \frac{S_1 \pm iS_2}{2}$$

Ahora, usando las relaciones de conmutación entre los elementos de la base  $\{S_1, S_2, S_3\}$ , obtenemos las relaciones de conmutación entre los elementos de la base esférica

$$[S_+, S_-] = \left[ \frac{S_1 + iS_2}{2}, \frac{S_1 - iS_2}{2} \right] = \frac{S_3}{2}$$

$$[S_+, S_-] = \frac{S_3}{2}$$

$$[S_3, S_{\pm}] = \left[ S_3, \frac{S_1 \pm iS_2}{2} \right] = \pm \frac{S_1 \pm iS_2}{2}$$

$$[S_3, S_{\pm}] = \pm S_{\pm}$$

¿Otras álgebras poseerán similar estructura?, es decir, si para otras álgebras existirá una base constituida por dos tipos de vectores similares a los vectores  $S_{\pm}$  y  $S_3$ . Veremos que para las álgebras de Lie denominadas semisimples existe dicha base (denominada base de Cartan-Weyl), para dar a conocer este notable resultado primeramente daremos el concepto de raíz y posteriormente hablaremos de raíces simples las cuales se muestran gráficamente en los denominados diagramas de Dynkin.



## Álgebra de Lie

Un álgebra de Lie  $L = \{V, [, ]\}$  es un espacio vectorial  $V$  (real o complejo) sobre el cual se define la operación binaria cerrada  $[, ]$  llamada conmutador (es decir, a todo par  $x, y \in V$ , le corresponde un elemento  $z \in V$  tal que  $z = [x, y]$ ). Dicha operación debe satisfacer condiciones de linealidad, antisimetría y la identidad de Jacobi

- 1)  $[\alpha x + \beta y, z] = \alpha[x, z] + \beta[y, z]$
- 2)  $[x, y] = -[y, x]$
- 3)  $[x, [y, z]] + [y, [z, x]] + [z, [x, y]] = 0$

donde  $x, y, z \in V$  y  $\alpha, \beta$  son números reales o complejos dependiendo del caso. Debe recalcar que la operación de conmutación para un álgebra de Lie arbitraria es una operación abstracta y no necesariamente es igual al conmutador del caso de los operadores. La dimensión del álgebra de Lie viene a ser la dimensión del espacio vectorial.

Por ejemplo, el espacio vectorial euclidiano  $L = \mathbb{R}^3$ , con la operación  $[\vec{x}, \vec{y}]$  definida como  $[\vec{x}, \vec{y}] \equiv \vec{x} \times \vec{y}$ , es un álgebra de Lie de dimensión 3.

Un álgebra de Lie se llama abeliano o conmutativo si  $[x, y] = 0$  para cualquier  $x, y \in L$ . Sean  $M, N$  dos subconjuntos de vectores del álgebra de Lie  $L$ , denotemos por  $[M, N]$  todos los vectores de la forma:  $[x, y]$ ,  $x \in M$ ,  $y \in N$ .

El subespacio  $N$  del álgebra de Lie  $L$  se llama subálgebra si  $[N, N] \subset N$  e ideal si  $[L, N] \subset N$ . Esta claro que el ideal automáticamente es un subálgebra. El  $\{0\}$  y toda el álgebra  $L$  son ideales de  $L$ , llamados ideales triviales.

Un subálgebra  $N$  que es abeliano e ideal se llama ideal abeliano.

Un álgebra de Lie se llama simple, si no posee ningún ideal a parte de los ideales triviales, y semisimple, si no posee ningún ideal abeliano.

Sea  $B = \{e_1, e_2, \dots, e_n\}$  una base del espacio vectorial  $L$  de dimensión  $n$ , por la linealidad del conmutador  $z = [x, y]$  y expresando los vectores  $x, y, z$  en la base  $B$  tenemos

$$z = [x, y]$$

$$\sum_{k=1}^n z^k e_k = \left[ \sum_{i=1}^n x^i e_i, \sum_{j=1}^n y^j e_j \right]$$

$$\sum_{k=1}^n z^k e_k = \sum_{i=1}^n \sum_{j=1}^n x^i y^j [e_i, e_j]$$

$$\sum_{k=1}^n z^k e_k = \sum_{k=1}^n \sum_{i=1}^n \sum_{j=1}^n x^i y^j C_{ij}^k e_k.$$

A partir de esta última ecuación deducimos que las componentes del vector  $z$  vienen dadas por

$$z^k = [x, y]^k = \sum_{i=1}^n \sum_{j=1}^n C_{ij}^k x^i y^j$$

donde

$$[e_i, e_j] = \sum_{k=1}^n C_{ij}^k e_k.$$

Los números  $C_{ij}^k$  se llaman constantes de estructura del álgebra de Lie  $L$ . Nótese que los valores de dichas constantes, dependen de la elección de la base. De los axiomas 2) y 3) se deduce

$$C_{ij}^k = -C_{ji}^k$$

$$\sum_{p=1}^n \left( C_{ij}^p C_{kp}^m + C_{jk}^p C_{ip}^m + C_{ki}^p C_{jp}^m \right) = 0$$

### **Criterio de semisimplicidad de Cartan**

En base a las constantes de estructura de un álgebra de Lie, definamos la siguiente cantidad simétrica (tensor métrico).

$$g_{ij} = g_{ji} = \sum_{k=1}^n \sum_{m=1}^n C_{ik}^m C_{jm}^k.$$

Un álgebra de Lie es semisimple si y solo si

$$\det(g_{ij}) \neq 0.$$

Este criterio nos permite determinar de manera práctica cuando un álgebra de Lie es semisimple o no.

A continuación mostraremos, usando el criterio de Cartan que el álgebra  $su(2)$  es semisimple.

### Semisimplicidad del álgebra $su(2)$

Las constantes de estructura del álgebra  $su(2)$  viene a ser el tensor antisimétrico  $i\epsilon_{ijk}$ . Por lo que para este caso el tensor métrico viene dado por

$$g_{ij} = \sum_{k,m=1}^3 -\epsilon_{imk} \epsilon_{jkm} = \sum_{k,m=1}^3 \epsilon_{imk} \epsilon_{jkm} = 2\delta_{ij}$$

entonces

$$\det(g_{ij}) = 8 \neq 0$$

por lo tanto queda mostrado que el álgebra  $su(2)$  es semisimple.

### Raíces y álgebras de Lie semisimples

Sea  $L$  un álgebra de Lie semisimple de dimensión  $n$ ; Cartan demostró que existe un conjunto  $\{H_i\}$  constituido por  $l$  vectores  $H_i \in L$   $\{i=1,2,\dots,l\}$  linealmente independientes tal que

$$[H_i, H_j] = 0. \tag{1}$$

El número entero  $l$  se denomina rango del álgebra, el rango del álgebra nos indica la máxima cantidad de vectores  $H_i$  que conmutan entre si. El conjunto  $\{H_i\} \subset L$  es un subálgebra de  $L$  denominada subálgebra de Cartan.

También puede deducirse que existe un conjunto  $\{E_\alpha\}$  constituido por  $n-l$  vectores  $E_\alpha \in L$  linealmente independientes tal que

$$[H_i, E_\alpha] = \alpha_i E_\alpha \quad (\alpha_i \in \mathbb{R}) \quad (2)$$

Al vector  $l$  dimensional  $\alpha = (\alpha_1, \alpha_2, \dots, \alpha_l) \neq \bar{0}$  se le denomina vector raíz no nulo o simplemente raíz no nula, existen  $n-l$  raíces no nulas distintas, llamaremos  $V_R$  al conjunto constituido por estas  $n-l$  raíces no nulas.

Los  $l$  vectores  $H_i$  y los  $n-l$  vectores  $E_\alpha$ , constituyen una base del álgebra  $L$ , dicha base se denomina base de Cartan-Weyl.

Por ejemplo la base de Cartan-Weyl del álgebra  $\mathfrak{su}(2)$  viene a ser el conjunto de vectores  $\{S_3, S_\pm\}$ , la subálgebra de Cartan para este caso vendría a ser el conjunto  $\{S_3\}$ . Denotemos la base de Cartan-Weyl como  $\{F_A\}$  ( $A=1, 2, \dots, l, \alpha, \beta, \gamma, \dots$ ) donde  $\alpha, \beta, \gamma, \dots \in V_R$ ,

$$F_i \equiv H_i \quad (i=1, 2, \dots, l) \quad (3)$$

$$F_\alpha \equiv E_\alpha \quad (\alpha \in V_R). \quad (4)$$

Las constantes de estructuras del álgebra  $C_{AB}^T$  vienen dadas a partir de las relaciones de conmutación entre los elementos de la base

$$[F_A, F_B] = \sum_T C_{AB}^T F_T. \quad (5)$$

Los índices con letras latinas mayúsculas tomarán valores  $1, 2, \dots, l, \alpha, \beta, \gamma, \dots$ ; los índices con letras latinas minúsculas tomarán valores de  $1, 2, \dots, l$ ; las letras griegas se emplearán para denotar a cualquiera de las  $n-l$  raíces no nulas.

Usando (1), (2), (3), (4), y (5) obtenemos

$$C_{ij}^T = 0 \quad (6)$$



$$C_{i\alpha}^j = 0 \quad (7)$$

$$C_{i\alpha}^\beta = \alpha_i \delta_{\alpha}^\beta. \quad (8)$$

A partir de la identidad de Jacobi,

$$[H_i, [E_\alpha, E_\beta]] + [E_\alpha, [E_\beta, H_i]] + [E_\beta, [H_i, E_\alpha]] = 0$$

obtenemos que si  $\alpha$  y  $\beta \in V_R$ ,

$$[H_i, [E_\alpha, E_\beta]] = (\alpha_i + \beta_i) [E_\alpha, E_\beta]$$

De esta última relación se demuestra que [2]

$$[E_\alpha, E_\beta] = \sum_{i=1}^l C_{\alpha\beta}^i H_i = \sum_{i=1}^l C_{\alpha, -\alpha}^i H_i \quad \text{si } \alpha + \beta = 0 \quad (9)$$

$$[E_\alpha, E_\beta] = N_{\alpha\beta} E_{\alpha+\beta} \quad \text{si } \alpha + \beta \in V_R \quad (10)$$

$$[E_\alpha, E_\beta] = 0 \quad \text{si } \alpha + \beta \neq 0 \text{ y } \alpha + \beta \notin V_R \quad (11)$$

donde  $N_{\alpha,\beta}$  es un coeficiente de proporcionalidad cuyo valor vendrá dado mas adelante. Las componentes del tensor métrico vienen definidos como

$$g_{AB} = \sum_{T,P} C_{AT}^P C_{BT}^T.$$

Ahora obtengamos las componentes  $g_{ij}$  de dicho tensor

$$\begin{aligned} g_{ij} &= \sum_{T,P} C_{iT}^P C_{jP}^T = \sum_{k=1}^l \sum_T C_{iT}^k C_{jk}^T + \sum_{\alpha} \sum_T C_{iT}^\alpha C_{j\alpha}^T \\ g_{ij} &= \sum_{k=1}^l \sum_T C_{iT}^k C_{jk}^T + \sum_{\alpha} \sum_{k=1}^l C_{ik}^\alpha C_{j\alpha}^k + \sum_{\alpha} \sum_{\beta} C_{i\beta}^\alpha C_{j\alpha}^\beta. \end{aligned} \quad (12)$$

Usando (6), (7), y (8) en 12 obtenemos

$$g_{ij} = \sum_{\alpha} \alpha_i \alpha_j$$

escribiendo esta última expresión en forma matricial

$$g = \sum_{\alpha} \alpha^T \alpha \quad (13)$$

donde la suma se extiende sobre todas las raíces no nulas.

En base a las componentes  $g_{ij}$  del tensor métrico, definimos el producto escalar en el espacio de los vectores raíces 1-dimensional como

$$(\alpha \cdot \beta) \equiv \sum_{i=1}^l \sum_{j=1}^l g^{ij} \alpha_i \beta_j.$$

Las componentes  $g^{ij}$  son definidas en base a las componentes  $g_{ij}$  del tensor métrico, de modo tal que

$$\sum_{k=1}^l g^{ik} g_{ki} = \delta_j^i.$$

Las componentes contravariantes de los vectores raíces se definen como

$$\alpha^i \equiv \sum_{j=1}^l g^{ij} \alpha_j.$$

De este modo el producto escalar se escribe

$$(\alpha \cdot \beta) \equiv \sum_{i=1}^l \alpha_i \beta^i \equiv \sum_{i=1}^l \alpha^i \beta_i = \sum_{i=1}^l \sum_{j=1}^l g_{ij} \alpha^i \beta^j$$

La longitud de una raíz  $\alpha$  se define como  $(\alpha, \alpha)^{1/2}$ .

A partir de la ecuación (13) se deduce que la matriz  $\{g_{ij}\}$  es una matriz simétrica real definida positiva, en consecuencia dicha matriz tendrá todos sus valores propios positivos. Una matriz cuadrada real  $A$  se llama definida positiva, si

$$\bar{x}^T A \bar{x} \geq 0 \text{ para todo } \bar{x} \in \mathbb{R}^n \text{ y } \bar{x}^T A \bar{x} = 0 \Leftrightarrow \bar{x} = 0.$$

A continuación mencionamos las propiedades generales de las raíces

(a1) Si  $\alpha$  es una raíz no nula, entonces  $k\alpha$  es una raíz si y solo si  $k = \pm 1$ .

Esto quiere decir que las raíces no nulas siempre aparecen en pares, se deduce entonces que  $n-1$  es un número entero par.

(a2) Si  $\alpha$  y  $\beta$  son dos raíces no nulas, existen dos números enteros no negativos  $p$  y  $q$  tal que

$$\beta - p\alpha, \beta - (p-1)\alpha, \dots, \beta, \beta + \alpha, \dots, \beta + q\alpha \quad (14)$$

son las únicas raíces del tipo  $\beta + k\alpha$ .

El conjunto de raíces (14) es llamado la  $\alpha$  cadena que contiene a  $\beta$ .

(a3) Si  $\alpha$  y  $\beta$  son dos raíces no nulas, entonces

$$\frac{2(\alpha \cdot \beta)}{(\alpha \cdot \alpha)} = p - q \quad (15)$$

y

$$\beta - (p - q)\alpha \quad (16)$$

es una raíz.

Si en (a2) y (a3) intercambiamos la raíz  $\alpha$  por  $\beta$  y  $\beta$  por  $\alpha$ , obtenemos relaciones similares pero con otros números enteros positivos  $p'$  y  $q'$ ; así por ejemplo si en la ecuación (15) intercambiamos  $\alpha$  con  $\beta$  tenemos

$$\frac{2(\alpha \cdot \beta)}{(\beta \cdot \beta)} = p' - q'. \quad (17)$$

Definimos el ángulo  $\varphi$  entre dos raíces no nulas  $\alpha$  y  $\beta$  como

$$\cos \varphi \equiv \frac{(\alpha \cdot \beta)}{[(\alpha \cdot \alpha)(\beta \cdot \beta)]^{1/2}}. \quad (18)$$

Elevando al cuadrado ambos miembros de la ecuación (18) y luego reemplazando (15) y (17) en dicha ecuación tenemos

$$\cos^2 \varphi = \frac{1}{4}(p-q)(p'-q'). \quad (19)$$

A partir de la ecuación (19), se ve que el ángulo  $\varphi$  puede asumir solo los siguientes valores

$$\varphi = 0^\circ \text{ (180}^\circ\text{)} \quad 30^\circ \text{ (150}^\circ\text{)} \quad 45^\circ \text{ (135}^\circ\text{)} \quad 60^\circ \text{ (120}^\circ\text{)} \quad 90^\circ. \quad (20)$$

A continuación se muestra un resumen de las relaciones de conmutación entre los elementos de la base de Cartan-Weyl  $\{H_i, E_\alpha\}$

$$[H_i, H_j] = 0 \quad (21)$$

$$[H_i, E_\alpha] = \alpha_i E_\alpha \quad (22)$$

$$[E_\alpha, E_{-\alpha}] = \sum_{i=1}^l \alpha^i H_i \quad \text{si } \alpha + \beta = 0$$

$$[E_\alpha, E_\beta] = N_{\alpha\beta} E_{\alpha+\beta} \quad \text{si } \alpha + \beta \in V_R \quad (23)$$

$$[E_\alpha, E_\beta] = 0 \quad \text{si } \alpha + \beta \in V_R \text{ y } \alpha + \beta \notin V_R$$

el coeficiente  $N_{\alpha\beta}$  viene dado por

$$N_{\alpha\beta} = \pm \left[ \frac{1}{2}(p+1)q(\alpha \cdot \alpha) \right]^{1/2} \quad (24)$$

donde  $p$  y  $q$  son dos números enteros positivos que caracterizan la  $\alpha$  cadena que contiene a  $\beta$ . Además dicho coeficiente satisface [2]

$$N_{\alpha\beta} = -N_{\beta\alpha} = -N_{-\alpha, -\beta} = N_{-\alpha, \alpha+\beta} = N_{\beta, -\alpha-\beta} \quad (25)$$



## Raíces simples

Una raíz no nula  $\alpha=(\alpha_1, \alpha_2, \dots, \alpha_l)$  se llama positiva si su primera componente no nula  $\alpha_i$  es un número positivo.

Por ejemplo, considerando las siguientes raíces no nulas

$$(1,0) \quad (1,1) \quad (1,-1) \quad (0,1) \quad (0,-1) \quad (-1,0) \quad (-1,1) \quad (-1,-1)$$

las primeras cuatro raíces de este conjunto son raíces positivas.

Una raíz no nula se llama simple si dicha raíz es positiva y no puede ser escrita como la suma de dos raíces positivas.

Del ejemplo anterior tenemos

$$(1,0) = (1,-1) + (0,1) \quad \text{y} \quad (1,1) = (1,0) + (0,1)$$

por lo tanto  $(1,0)$  y  $(1,1)$  no son raíces simples. Las raíces  $(0,1)$  y  $(1,-1)$  sí son raíces simples.

Un álgebra de Lie semisimple de rango  $l$  posee  $l$  raíces simples, es decir el número de raíces simples es igual al rango del álgebra.

A continuación mencionamos las propiedades más importantes de las raíces simples

(c1) Si  $\alpha$  y  $\beta$  son dos raíces simples,  $\alpha - \beta$  no es una raíz y  $(\alpha, \beta) \leq 0$ .

(c2) Cualquier raíz positiva puede ser escrita como una combinación lineal de raíces simples con coeficientes enteros positivos

$$\alpha_{\text{positivo}} = \sum k_i \alpha_{\text{simple}}^{(i)} \quad (k_i \geq 0 \text{ entero}) \quad (26)$$

(c3) Si  $\alpha$  y  $\beta$  son dos raíces simples, el ángulo entre ellos  $\varphi_{\alpha\beta}$  puede tomar los valores de  $90^\circ$ ,  $120^\circ$ ,  $135^\circ$  o  $150^\circ$ .

Si  $(\alpha, \alpha) \leq (\beta, \beta)$ , entonces

$$\frac{(\beta \cdot \beta)}{(\alpha \cdot \alpha)} = \begin{cases} 1 & \text{para } \varphi_{\alpha\beta} = 120^\circ \\ 2 & \text{para } \varphi_{\alpha\beta} = 135^\circ \\ 3 & \text{para } \varphi_{\alpha\beta} = 150^\circ \\ \text{in determinado} & \text{para } \varphi_{\alpha\beta} = 90^\circ \end{cases} \quad (27)$$

(c4) Si  $\alpha$  es una raíz positiva no simple, es posible encontrar una raíz simple  $\alpha^{(k)}$  tal que  $\alpha - \alpha^{(k)}$  es otra raíz positiva.

Las propiedades de las raíces simples, junto con las propiedades generales de las raíces, permite determinar las demás raíces a partir de las raíces simples.

Con el fin de obtener todas las raíces a partir de las raíces simples, lo primero que haremos será dar el concepto de nivel.

Si  $\alpha$  es una raíz positiva, de acuerdo a la ecuación (26), existe  $l$  números enteros positivos  $k_i$  ( $i=1,2,\dots,l$ ) que permiten expresar dicha raíz como una combinación lineal de raíces simples. Si

$$\sum_{i=1}^l k_i = N, \quad (28)$$

decimos que la raíz positiva  $\alpha$  pertenece al  $N$ ésimo nivel. En particular, todas las raíces simples pertenecen al primer nivel ( $N=1$ ). De la propiedad (c4) vemos que una raíz positiva del  $N$ ésimo nivel, puede ser obtenida por la suma de una raíz simple con alguna raíz positiva del  $(N-1)$ ésimo nivel. En particular si un nivel está vacío, es decir si dicho nivel no contiene ninguna raíz positiva, todos los niveles sucesivos también estarán vacíos.

Para hallar todas las raíces, primero debemos hallar todas las raíces positivas, una vez obtenidas todas estas raíces (raíces positivas), las demás raíces se obtendrán usando la propiedad (a1).

Ahora veremos los pasos a seguir para obtener todas las raíces positivas.

Asumimos que conocemos todas las raíces positivas hasta un cierto nivel,

digamos hasta el  $N$ ésimo nivel.

Si  $\alpha = \sum k_i \alpha^{(i)}$  es una raíz positiva que pertenece a este nivel ( $N$ ésimo), consideramos la  $\alpha^{(k)}$  cadena ( $\alpha^{(k)}$  es una raíz simple) que contiene a  $\alpha$ .

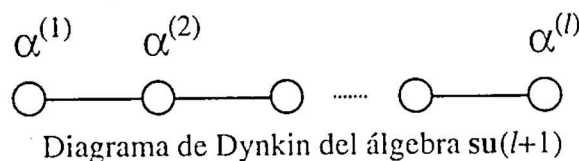
De acuerdo a la ecuación (15) tenemos,

$$p - q = \frac{2\alpha \cdot \alpha^{(k)}}{\alpha^{(k)} \cdot \alpha^{(k)}} = \sum_{i=1}^l k_i \frac{2\alpha^{(i)} \cdot \alpha^{(k)}}{\alpha^{(k)} \cdot \alpha^{(k)}} \quad (\alpha^{(i)}, \alpha^{(k)} \text{ raíces simples}) \quad (29)$$

Como conocemos todas las raíces positivas hasta el nivel  $N$ ésimo, podemos verificar si  $\alpha - m\alpha^{(k)}$  es o no una raíz para algún número entero positivo  $m$ , de esta manera podemos conocer el número  $p$  relativo a la  $\alpha^{(k)}$  cadena que contiene a  $\alpha$ , luego usando la ecuación (29) determinamos el valor de  $q$ , y si  $q \geq 1$ , entonces  $\alpha + \alpha^{(k)}$  es una raíz del  $(N+1)$ ésimo nivel. De esta manera podemos obtener todas las raíces positivas a partir del primer nivel, vale decir a partir de las raíces simples por medio de un procedimiento recurrente.

## Diagrama de Dynkin

Los diagramas de Dynkin muestran de manera abreviada todas las raíces simples correspondientes a cada álgebra de Lie semisimple. En un diagrama de Dynkin cada raíz simple es representada por un pequeño círculo. Pares de raíces simples (círculos) que están conectados por una, dos, o tres líneas, corresponden a ángulos de  $120^\circ$ ,  $135^\circ$ , o  $150^\circ$  entre dichas raíces. Pares de círculos que no están conectados, corresponden a pares de raíces simples cuyo ángulo entre ellos es de  $90^\circ$ . Los círculos negros corresponden a las raíces simples con menor longitud respecto de las longitudes de las demás raíces simples, mientras que los círculos blancos corresponden a las raíces simples con mayor longitud respecto de las longitudes de las demás raíces simples. Si todos los círculos son de color blanco, quiere decir que todas las raíces simples tienen la misma longitud. Como ejemplo a continuación mostramos el diagrama de Dynkin del álgebra  $\mathfrak{su}(l+1)$  ( $l=1,2,3,\dots$ )



## Ejemplos

### Álgebra de Lie semisimple de rango 1

Un álgebra de Lie de rango uno poseerá solo una raíz simple, denominemos  $\alpha^{(1)}$  a dicha raíz. Para este caso el primer nivel  $N=1$  contiene a la raíz  $\alpha^{(1)}$ . De acuerdo con la ecuación (26) el segundo nivel  $N=2$ , puede contener solo a la raíz  $2\alpha^{(1)}$ , pero de acuerdo con la propiedad (a1)  $-\alpha^{(1)}$  y  $\alpha^{(1)}$  son las únicas raíces no nulas del tipo  $k\alpha^{(1)}$ , por lo tanto  $2\alpha^{(1)}$  no es una raíz, es decir el segundo nivel está vacío, entonces tenemos solo una raíz positiva  $\alpha^{(1)}$ , en consecuencia todas las raíces no nulas vienen a ser  $\pm\alpha^{(1)}$ . Puesto que tenemos 2 raíces no nulas y el rango del álgebra es uno, la dimensión del álgebra es para este caso

$$n = 2 + 1 = 3.$$

Si escogemos  $\alpha^{(1)}=1$ , usando (13) obtenemos  $g_{11}=2$ , a partir de esta última ecuación tenemos  $g^{11}=1/2$ .

Usamos (22) y (23) obtenemos

$$\begin{aligned} [H_1, E_1] &= E_1 \\ [H_1, E_{-1}] &= -E_{-1} \\ [E_1, E_{-1}] &= \frac{1}{2} H_1. \end{aligned} \tag{30}$$

Sea  $\{S_1, S_2, S_3\}$ , una base de esta álgebra de Lie de dimensión 3, cuya relación con la base de Cartan-Weyl  $\{H_1, E_1, E_{-1}\}$  viene dada por

$$\begin{aligned} S_1 &= E_1 + E_{-1} \\ S_2 &= i(E_{-1} - E_1) \\ S_3 &= H_1 \end{aligned} \tag{31}$$

Empleando (30) y (31), obtenemos

$$[S_i, S_j] = \sum_{k=1}^3 i\epsilon_{ijk} S_k$$



Vemos entonces que esta álgebra de dimensión 3 de rango uno, es el álgebra  $su(2)$ .

## Álgebras de Lie semisimples de rango 2

Existen varios ejemplos de álgebras de Lie semisimples de rango dos, dicha álgebras poseerán 2 raíces simples, llamando  $\alpha^{(1)}$  y  $\alpha^{(2)}$  a dichas raíces. En esta parte discutiremos en detalle el caso cuando el ángulo entre las dos raíces simples es de  $\varphi_{\alpha^{(1)} \alpha^{(2)}} = 120^\circ$ . De acuerdo a (c3) tenemos  $(\alpha^{(1)} \cdot \alpha^{(1)})/(\alpha^{(2)} \cdot \alpha^{(2)})=1$ . Tomando  $\alpha^{(2)} \cdot \alpha^{(2)}=1$  tendremos que  $\alpha^{(1)} \cdot \alpha^{(1)}=1$  y  $\alpha^{(1)} \cdot \alpha^{(2)}=-1/2$ .

El primer nivel  $N=1$  contiene las raíces  $\alpha^{(1)}$  y  $\alpha^{(2)}$ . El segundo nivel  $N=2$  puede contener las raíces  $2\alpha^{(1)}$ ,  $2\alpha^{(2)}$  o  $\alpha^{(1)}+\alpha^{(2)}$ . Sin embargo, debido a la propiedad (a1)  $2\alpha^{(1)}$  y  $2\alpha^{(2)}$  no pueden ser raíces. Ahora verificaremos si  $\alpha^{(1)}+\alpha^{(2)}$  es o no una raíz, para ello consideremos la  $\alpha^{(2)}$  cadena que contiene a  $\alpha^{(1)}$ . De acuerdo a (c1),  $\alpha^{(1)}-\alpha^{(2)}$  no es una raíz, por lo tanto  $p=0$ . De la ecuación (29) tenemos

$$-q = \frac{2\alpha^{(1)} \cdot \alpha^{(2)}}{\alpha^{(2)} \cdot \alpha^{(2)}} = -1$$

por lo tanto  $q=1$  y entonces  $\alpha^{(1)}+\alpha^{(2)}$  es una raíz positiva del segundo nivel. Para el tercer nivel, sabemos que todas las raíces de este nivel son obtenidas mediante la suma de una de las raíces del segundo nivel (cualquiera de ellas) con alguna raíz simple.

Consideremos ahora la  $\alpha^{(2)}$  cadena que contiene a  $\alpha^{(1)}+\alpha^{(2)}$  (raíz del segundo nivel). Dado que  $(\alpha^{(1)}+\alpha^{(2)})-\alpha^{(2)}=\alpha^{(1)}$  es una raíz pero  $(\alpha^{(1)}+\alpha^{(2)})-2\alpha^{(2)}=\alpha^{(1)}-\alpha^{(2)}$  no es una raíz, tenemos que  $p=1$ . De la ecuación (29) tenemos

$$1-q = \frac{2\alpha^{(2)} \cdot (\alpha^{(1)} + \alpha^{(2)})}{\alpha^{(2)} \cdot \alpha^{(2)}} = \frac{2[\alpha^{(1)} \cdot \alpha^{(2)} + \alpha^{(2)} \cdot \alpha^{(2)}]}{\alpha^{(2)} \cdot \alpha^{(2)}} = -1 + 2 = 1$$

por lo tanto  $q=0$  y entonces  $(\alpha^{(1)}+\alpha^{(2)})+\alpha^{(2)}=\alpha^{(1)}+2\alpha^{(2)}$  no es una raíz pues  $q=0$ , si  $q$  hubiese sido igual a 1 ahí si  $\alpha^{(1)}+2\alpha^{(2)}$  hubiera sido una raíz positiva del tercer nivel.

Procediendo de manera análoga al caso anterior pero ahora considerando la  $\alpha^{(1)}$  cadena que contiene a  $\alpha^{(1)}+\alpha^{(2)}$ , se demuestra que  $2\alpha^{(1)}+\alpha^{(2)}$  no es una raíz. Por

lo tanto el tercer nivel  $N=3$  esta vacío y entonces tenemos solo las tres raíces positivas,

$$\alpha^{(1)}, \alpha^{(2)}, \alpha^{(1)} + \alpha^{(2)}.$$

En consecuencia todas las raíces no nulas vienen a ser,

$$\pm \alpha^{(1)}, \pm \alpha^{(2)}, \text{ y } \pm (\alpha^{(1)} + \alpha^{(2)}).$$

Debido a que tenemos 6 raíces no nulas y el rango del álgebra es dos, la dimensión del álgebra es en este caso

$$n = 6 + 2 = 8.$$

Si elegimos  $\alpha^{(1)} = \left(\frac{1}{2\sqrt{3}}, \frac{1}{2}\right)$  y  $\alpha^{(2)} = \left(\frac{1}{2\sqrt{3}}, -\frac{1}{2}\right)$ , obtenemos las demas raíces no nulas  $-\alpha^{(1)} = \left(-\frac{1}{2\sqrt{3}}, -\frac{1}{2}\right)$ ,  $\alpha^{(2)} = \left(-\frac{1}{2\sqrt{3}}, \frac{1}{2}\right)$ ,  $\alpha^{(1)} + \alpha^{(2)} = \left(\frac{1}{\sqrt{3}}, 0\right)$  y  $-(\alpha^{(1)} + \alpha^{(2)}) = \left(-\frac{1}{\sqrt{3}}, 0\right)$ , luego usando (13) obtenemos  $g_{ij} = \delta_{ij}$ , a partir de esta última ecuación tenemos  $g^{ij} = \delta_{ij}$ .

Para este caso  $\{H_1, H_2, E(\frac{1}{2\sqrt{3}}, \frac{1}{2}), E(\frac{1}{2\sqrt{3}}, -\frac{1}{2}), E(-\frac{1}{2\sqrt{3}}, -\frac{1}{2}), E(-\frac{1}{2\sqrt{3}}, \frac{1}{2}), E(\frac{1}{\sqrt{3}}, 0), E(-\frac{1}{\sqrt{3}}, 0)\}$ , viene a ser la base de Cartan-Weyl, usando (22), (23), (24) y (25), obtenemos las relaciones de conmutación entre los elementos de esta base

$$[H_1, E(\pm\frac{1}{2\sqrt{3}}, \pm\frac{1}{2})] = \pm \frac{1}{2\sqrt{3}} E(\pm\frac{1}{2\sqrt{3}}, \pm\frac{1}{2})$$

$$[H_1, E(\pm\frac{1}{2\sqrt{3}}, \mp\frac{1}{2})] = \pm \frac{1}{2\sqrt{3}} E(\pm\frac{1}{2\sqrt{3}}, \mp\frac{1}{2})$$

$$[H_1, E(\pm\frac{1}{\sqrt{3}}, 0)] = \pm \frac{1}{\sqrt{3}} E(\pm\frac{1}{\sqrt{3}}, 0)$$

$$[H_2, E(\pm \frac{1}{2\sqrt{3}}, \pm \frac{1}{2})] = \pm \frac{1}{2} E(\pm \frac{1}{2\sqrt{3}}, \pm \frac{1}{2})$$

$$[H_2, E(\pm \frac{1}{2\sqrt{3}}, \mp \frac{1}{2})] = \mp \frac{1}{2} E(\pm \frac{1}{2\sqrt{3}}, \mp \frac{1}{2})$$

$$[H_2, E(\pm \frac{1}{\sqrt{3}}, 0)] = 0$$

$$[E(\frac{1}{2\sqrt{3}}, \frac{1}{2}), E(-\frac{1}{2\sqrt{3}}, -\frac{1}{2})] = \frac{1}{2\sqrt{3}} H_1 + \frac{1}{2} H_2$$

$$[E(\frac{1}{2\sqrt{3}}, -\frac{1}{2}), E(-\frac{1}{2\sqrt{3}}, \frac{1}{2})] = \frac{1}{2\sqrt{3}} H_1 - \frac{1}{2} H_2$$

$$[E(\frac{1}{\sqrt{3}}, 0), E(-\frac{1}{\sqrt{3}}, 0)] = \frac{1}{\sqrt{3}} H_1$$

$$[E(\frac{1}{2\sqrt{3}}, \frac{1}{2}), E(\frac{1}{2\sqrt{3}}, -\frac{1}{2})] = \frac{1}{\sqrt{6}} E(\frac{1}{\sqrt{3}}, 0)$$

$$[E(\frac{1}{2\sqrt{3}}, \frac{1}{2}), E(-\frac{1}{\sqrt{3}}, 0)] = -\frac{1}{\sqrt{6}} E(-\frac{1}{2\sqrt{3}}, \frac{1}{2})$$

$$[E(\frac{1}{2\sqrt{3}}, -\frac{1}{2}), E(-\frac{1}{\sqrt{3}}, 0)] = \frac{1}{\sqrt{6}} E(-\frac{1}{2\sqrt{3}}, -\frac{1}{2})$$

Esta álgebra de dimensión 8 de rango dos, es el álgebra  $su(3)$  [2].

## Conclusiones

En este trabajo damos una regla práctica para obtener todas las raíces de una álgebra de Lie semisimple a partir de las raíces simples. Una vez determinadas todas las raíces (como se vió en los ejemplos), se puede hallar las relaciones de conmutación entre los elementos de la base de Cartan-Weyl.

Un álgebra de Lie posee varias bases, vimos que para un álgebra de Lie semisimple existe una base especial denominada base de Cartan Weyl, donde las relaciones de conmutación entre los elementos de esta base adoptan una forma que es muy usada en Física.

## BIBLIOGRAFÍA

- [1] *Greiner, Walter, Muller, Berndt, Quantum Mechanics: Symmetries, Berlin, Springer-Verg 1994.*
- [2] *L. Fonda, G. C. Ghirardi, Symmetry Principles in Quantum Physics, Textbook Binding, 1986.*

# El método de Newton (amortiguado) para desigualdades variacionales

---

*Hermes Pantoja Carhuavilca\*,  
William Echegaray Castillo\*\**

## *RESUMEN*

El trabajo presenta un algoritmo de Newton (Amortiguado) para resolver problemas de desigualdad variacional basando la formulación de estos problemas como un sistema de ecuaciones usando la aplicación Minty. El propósito de este método es asegurar la convergencia y una convergencia cuadrática local bajo la suposición de regularidad. Bajo la suposición de regularidad débil y algunas condiciones mild, el algoritmo modificado demuestra que siempre existe una dirección descendente y converge a la solución.

## *ABSTRACT*

The work presents Newton algorithm (Muffled) to solve problems of inequality variacional basing the formulation of these problems like a system of equations using the application Minty. The purpose of this method is to assure the convergence and a convergence quadratic local under the supposition of regularity. Under the supposition of weak regularity and some conditions mild, the modified algorithm demonstrates that a descending address always exists and it converges to the solution.

---

\* *Fac. de Ciencias, UNI.*

\*\* *Asesor.*

## 1. Formulación de la ecuación del problema de desigualdad variacional

En este capítulo definimos los problemas de desigualdad variacional, complementariedad lineal y no lineal; establecemos sus relaciones. Además presentamos la definición de una función B-diferenciable y varias propiedades importantes que se derivan de dicha función, formulamos los problemas como un sistema de ecuaciones usando la aplicación Minty. Finalizamos este capítulo introduciendo una función mérito que permitirá desarrollar un algoritmo convergente en la búsqueda de la solución al sistema de ecuaciones.

### 1.1 Las definiciones del problema y hechos básicos

#### Definición 1.1.1 (Problema de desigualdad variacional)

Sea  $X$  un subconjunto no vacío de  $\mathbb{R}^n$  y  $F: \mathbb{R}^n \rightarrow \mathbb{R}^n$  una aplicación. El problema de desigualdad variacional, denotado por  $VI(X, F)$ , es encontrar un  $x^* \in X$  tal que:

$$F(x^*)^T(x-x^*) \geq 0, \quad \forall \quad x^* \in X \quad (1.1)$$

Uno típicamente asume que el conjunto  $X$  es cerrado y convexo.

#### Definición 1.1.2 (Problema de complementariedad no lineal)

Sea  $F: \mathbb{R}^n \rightarrow \mathbb{R}^n$  una aplicación. El problema de complementariedad no lineal, denotado por  $NCP(F)$ , es encontrar un vector  $x^* \in \mathbb{R}_+^n$  tal que:

$$F(x^*)^T(x^*) = 0, \quad F(x^*) \in \mathbb{R}_+^n \quad (1.2)$$

Cuando  $F$  es una función afín de  $x$ , decimos  $F(x) = q + Mx$ , donde  $q \in \mathbb{R}^n$  es un vector dado y  $M \in \mathbb{R}^{n \times n}$  es una matriz dada entonces, el problema  $NCP(F)$  se reduce al Problema de Complementariedad Lineal, lo cual denotaremos por  $LCP(q, M)$ .

**Definición 1.1.3 (Problema de complementariedad lineal)**

Sea  $F: \mathbb{R}^n \rightarrow \mathbb{R}^n$  una aplicación. El problema de complementariedad lineal, denotado como LCP ( $q, M$ ), es encontrar un vector  $x^* \in \mathbb{R}_+^n$  tal que:

$$q + Mx^* \geq 0, \quad x^* \geq 0 \quad (1.3)$$

$$(q + Mx^*)^T x^* \geq 0 \quad (1.4)$$

donde:  $q \in \mathbb{R}^n$  es un vector dado y  $M \in \mathbb{R}^{n \times n}$  es una matriz dada.

**Ejemplo 1.**

Sea  $f: I = [a, b] \rightarrow \mathbb{R}$  una función continuamente diferenciable. Buscamos los puntos  $x_0 \in I$  tal que:

$$f(x_0) = \min_{x \in I} f(x)$$

Tres casos pueden suceder:

1. Si  $a < x_0 < b$ , entonces  $f'(x_0) = 0$ ;
2. Si  $x_0 = a$ , entonces  $f'(x_0) \geq 0$ ; y
3. Si  $x_0 = b$ , entonces  $f'(x_0) \leq 0$

Lo que es equivalente a:  $f'(x_0)(x - x_0) \geq 0 \quad \forall x \in I$ . Esta desigualdad también es una Desigualdad Variacional.

**1.2 Formulación del problema**

Para el problema de desigualdad variacional VI ( $X, F$ ) donde:

$$X = \{x \in \mathbb{R}^n / g_i(x) \leq 0, i=1, \dots, m; h_j(x)=0, j=1, \dots, p\}$$

$F, g, h$ , son continuamente diferenciables, si  $x^*$  es solución del problema y si una

calificación de restricción es satisfecha en  $x^*$  entonces existen vectores  $y^* \in \mathbb{R}^m$  y  $v^* \in \mathbb{R}^p$  tal que las siguientes condiciones de complementariedad mixtas son satisfecha en  $(x^*, y^*, v^*)$

$$F(x) + \nabla g(x)y + \nabla h(x)v = 0 \quad (1.5)$$

$$y \geq 0, \quad g(x) \leq 0, \quad g(x)^T y = 0 \quad (1.6)$$

$$h(x) = 0 \quad (1.7)$$

Para escribir (1.5) – (1.7) como un sistema de ecuaciones, definimos:

$$u_i^+ = \max\{u_i, 0\}, \quad u_i^- = \min\{u_i, 0\}, \quad i=1 \dots m;$$

$$u^+ = (u_1^+, \dots, u_m^+)^T, \quad u^- = (u_1^-, \dots, u_m^-)^T$$

Debido a que  $y \geq 0$ , sea  $u^+ = y$ . Debido a que  $g(x) \leq 0$  y  $g(x)^T y = 0$ , podemos escribir que  $u^- = g(x)$ . Así podemos verificar que (1.5) – (1.7) es equivalente al siguiente sistema de ecuaciones no lineales:

$$F(x) + \nabla g(x)u^+ + \nabla h(x)v = 0 \quad (1.8)$$

$$g(x) - u^- = 0 \quad (1.9)$$

$$h(x) = 0 \quad (1.10)$$

En el contexto del problema de complementariedad no-lineal donde  $g(x) = -x$  el sistema de ecuaciones [08] – [10] se convierten en:

$$F(x) - u^+ = 0 \quad (1.11)$$

$$x + u^- = 0 \quad (1.12)$$

Puede además ser reducido a:

$$F(x^+) - x^- = 0 \quad (1.13)$$

donde:

$$x^+ = (x_1^+, \dots, x_m^+)^T, \quad x^- = (x_1^-, \dots, x_m^-)^T, \quad y$$

$$x_i^+ = \max(x_i, 0), \quad x_i^- = \min(x_i, 0), \quad i=1, \dots, m.$$



Las construcciones de  $x^+$  y  $x^-$  son frecuentemente referidos como la aplicación Minty.

Sea  $H: \mathbb{R}^{n+m+p} \rightarrow \mathbb{R}^{n+m+p}$  la función definida por:

$$H(z) = \begin{pmatrix} F(x) + \nabla g(x)u^+ + \nabla h(x)v \\ -g(x) + u^- \\ -h(x) \end{pmatrix} \quad (1.14)$$

donde  $z = (x, u, v)^T \in \mathbb{R}^{n+m+p}$ . Usando esta función, el problema para resolver VI  $(X, F)$  se convierte en el problema para resolver el siguiente sistema de ecuaciones:

$$H(z) = 0 \quad (1.15)$$

Para encontrar un cero de la función  $H$ , el método ordinario de Newton no puede ser aplicado ya que  $H$  no es una función continuamente diferenciable. El objetivo básico de este trabajo es desarrollar un método generalizado de Newton para las funciones los cuales no son  $F$  (réchet)-diferenciable pero que tienen  $B$ -derivadas en todas partes.

### 1.3 Las definiciones de derivadas

#### Definición 1.3.1 (Derivada de Fréchet)

Sean  $X$  e  $Y$  dos espacios normados y  $F$  es una aplicación Fréchet en  $x_0 \in X$  si y sólo si, existe una aplicación  $T \in L(X, Y)$ , tal que:

$$F(x_0+h) - F(x_0) = T(h) + o(\|h\|)$$

donde  $L(X, Y) = \{T: X \rightarrow Y / T \text{ es lineal y continua}\}$

Esto implica, en particular que:

$$\lim_{\|h\| \rightarrow 0} \frac{\|F(x_0 + h) - F(x_0) - T(h)\|}{\|h\|} = 0 \quad (1.16)$$

A la funcional  $T$  la llamaremos diferencial de Fréchet en  $x_0$  y la denotaremos por  $F'(x_0) := T$ .

Si  $X = \mathbb{R}^n$ , entonces la diferencial de Fréchet coincide con el concepto de diferencial del análisis clásico.

### Definición 1.3.2 (Derivada direccional)

Sea  $S \subseteq \mathbb{R}^n$  un conjunto no vacío y  $f: S \rightarrow \mathbb{R}$  una función continua. Sea  $z \in S$  y  $d \in \mathbb{R}^n$  un vector diferente de cero tal que  $z + \lambda d \in S$  para  $\lambda > 0$  suficientemente pequeño. La derivada direccional de  $f$  en  $z$  a lo largo del vector  $d$ , denotado por  $f'(z, d)$  es dado por el siguiente límite si este existe:

$$f'(z, d) = \lim_{\lambda \rightarrow 0^+} \frac{f(z + \lambda d) - f(z)}{\lambda}$$

La derivada direccional de una función vectorial  $H$ , definido en (1.14), de un vector  $z$  a lo largo de  $d$ , es denotado por  $H'(z, d)$ .

Un concepto cercanamente relacionado a la derivada dirección es la  $B$ -derivada

### Definición 1.1.3 (La $B$ -derivada)

Una función  $H: \mathbb{R}^n \rightarrow \mathbb{R}^m$  se dice que es  $B$ -diferenciable en un punto  $z$  si  $H$  es lipschitziana en una vecindad de  $z$  y además existe una función  $BH(z): \mathbb{R}^n \rightarrow \mathbb{R}^m$ , llamada la  $B$ -derivada de  $H$  en  $z$ , la cual es una función homogénea positiva de grado 1 (es decir:  $BH(z)(\lambda v) = \lambda BH(z)v$ ,  $v \in \mathbb{R}^n$ ,  $\lambda \geq 0$ ) tal que:

$$\lim_{v \rightarrow 0} \frac{H(z+v) - H(z) - BH(z)(v)}{\|v\|} = 0$$

Si  $H$  es  $B$ -diferenciable en todos los puntos de un conjunto  $S$ , entonces se dice que  $H$  es  $B$ -diferenciable en  $S$ .

### Definición 1.3.4

La  $B$ -derivada  $BH(z)$  se dice fuerte si:

$$\lim_{(v,w) \rightarrow (0,0)} \frac{H(z+v) - H(z+w) - BH(z)(v-w)}{\|v-w\|} = 0$$

### Teorema 1.3.1.

Sea  $H(z)$  definida por (1.14). Entonces:

- a)  $H$  es  $B$ -diferenciable en todas partes, y la  $B$ -derivada de  $H$  en un vector  $z = (x, y, v)^T$  a lo largo de la dirección  $d = (dz, du, dv)^T$ , denotado por  $H'(z, d)$ , ó  $BH(z) d$  está dado por:

$$\left( \begin{aligned} & \left[ \nabla F(x) + \sum_{i=0}^m u_i^+ \nabla^2 g_i(x) + \sum_{j=1}^p v_j \nabla^2 h_j(x) \right] dx + \nabla g(x) du^+ + \nabla h(x) dv \\ & - \nabla g(x)^T dx + du^- \\ & - \nabla h(x)^T dx \end{aligned} \right) \quad (1.17)$$

donde:

$$du_i^+ = \begin{cases} du_i & , \text{ Si } u_i > 0 \\ \max(du_i, 0) & , \text{ Si } u_i = 0 \\ 0 & , \text{ Si } u_i < 0 \end{cases} \quad y \quad du_i^- = \begin{cases} 0 & , \text{ Si } u_i > 0 \\ \min(du_i, 0) & , \text{ Si } u_i = 0 \\ du_i & , \text{ Si } u_i < 0 \end{cases}$$

b) Denotando:

$$\alpha(z) = \{i / u_i > 0\}$$

$$\beta(z) = \{i / u_i = 0\}$$

$$\lambda(z) = \{i / u_i < 0\}$$

Entonces  $H$  es  $F$ -diferenciable si y solo si el conjunto  $p$  es vacío. Además si el conjunto  $\beta$  es vacío entonces la  $B$ -derivada  $BH(z)$  es fuerte.

## 1.4 Introducción de una función mérito

Para desarrollar un algoritmo convergente en la búsqueda de la solución al problema [15], necesitamos introducir una función mérito  $\theta : \mathbb{R}^r \rightarrow \mathbb{R}$  definida por:

$$\theta(z) = \frac{1}{2} H(z)^T H(z) = \frac{1}{2} \|H(z)\|^2 \quad (1.18)$$

Aquí,  $r = n + m + p$ .

Claramente, encontrar un cero de  $H$  es equivalente a resolver un punto mínimo global  $z^*$  de  $\theta$  con  $\theta(z^*) = 0$ .

La función  $\theta$  es generalmente no  $F$ -diferenciable. Sin embargo, esto es siempre  $B$ -diferenciable.

### Teorema 1.4.1

Sea  $\theta : \mathbb{R}^r \rightarrow \mathbb{R}$  definido por (1.18). Entonces:

Si  $\beta(z) = \emptyset$  entonces  $\theta$  es fuertemente  $F$ -diferenciable en  $z$ . Además, si  $z^*$  es solución de  $H$  entonces  $\theta$  es fuertemente  $F$ -diferenciable en  $z^*$  y  $\nabla\theta(z^*) = 0$

Se nota que en un cero de  $H$ , el conjunto  $\beta$  puede no ser vacío, o que  $H$  puede no ser  $F$ -diferenciable. En otras palabras, el ser  $\beta(z^*) = \emptyset$  no es condición necesaria para la  $F$ -diferenciabilidad en  $\theta$  en  $z^*$ .

## 2. El algoritmo de Newton (amortiguado)

El algoritmo de Newton (amortiguado) fue usado por Harker y Pang para resolver problemas de complementariedad y por Harker y Xiao para resolver problemas de complementariedad no lineal cuyos experimentos numéricos sugieren que el método de Newton (amortiguado) es generalmente más eficiente y robusto que el método tradicional de Newton. Esta aproximación toma muy pocas iteraciones que permiten converger a la solución en donde el método tradicional falla. Describiremos el método de Newton (amortiguado) como sigue:

### *Paso 0:*

Sea  $z^0 \in \mathbb{R}^{n+m+p}$  un vector arbitrario inicial, y dados  $\mu$  y  $\sigma$  escalares donde  $\mu \in (0,1)$  y  $\sigma \in (0, \frac{1}{2})$ ,  $\varepsilon > 0$  pequeño.

Paso  $k$  ( $k=1,2,\dots$ ): Si  $\|H(z^k)\| \leq \varepsilon$  parar, en otro caso, generar  $z^{k+1}$  realizando los siguientes dos pasos:

1. Resolver la ecuación de Newton:

$$H(z^k) + H'(z^k)d^k = 0 \quad (2.1)$$

para la dirección  $d^k \in \mathbb{R}^{n+m+p}$

2.- Sea  $\lambda_k = \mu^{m_k}$  donde:

$$m_k = \text{Min} \{ \{ m \in \mathbb{Z}_+ \cup \{0\} / \theta(z^k) - \theta(z^k + \mu^m d^k) \geq 2\sigma \mu^m \theta(z^k) \} \} \quad (2.2)$$

Fijar  $z^{k+1} = z^k + \lambda_k d^k$ , e ir al paso  $k + 1$ .

Para analizar este algoritmo, definimos:

$$\mu_\alpha = \{ \mu_i / i \in \alpha(z) \},$$

$$\mu_\beta = \{ \mu_i / i \in \beta(z) \},$$

$$\mu_\gamma = \{ \mu_i / i \in \gamma(z) \},$$

$$g_\alpha(x) = \{ g_i(x) / i \in \alpha(z) \},$$

$$g_\beta(x) = \{ g_i(x) / i \in \beta(z) \},$$

$$g_\gamma(x) = \{g_i(x) / i \in \gamma(z)\},$$

$$H_F(z) = F(x) + \nabla g_\alpha(x) \mu_\alpha + \nabla h(x) v,$$

$$H_\alpha(z) = -g_\alpha(x),$$

$$H_\beta(z) = -g_\beta(x),$$

$$H_\gamma(z) = -g_\gamma(x) + \mu_\gamma,$$

$$H_h(z) = -h(x).$$

Usando esta notación, la ecuación (2.1) puede ser escrito como:

$$\begin{aligned} H_F(z) + [\nabla F(x) + \sum_{i \in \alpha(z)} \mu_i \nabla^2 g_i(x) + \sum_{j=1}^p v_j \nabla^2 h_j(x)] dx + \\ + \nabla g_\alpha(x) d\mu_\alpha + \nabla g_\beta(x) \max(0, du_\beta) + \nabla h(x) dv = 0 \end{aligned} \quad (2.3)$$

$$H_\alpha(z) - \nabla g_\alpha(x)^T dx = 0 \quad (2.4)$$

$$H_\beta(x) - \nabla g_\beta(x)^T dx + \min(0, d\mu_\beta) = 0 \quad (2.5)$$

$$H_\gamma(z) - \nabla g_\gamma(x)^T dx + d\mu_\gamma = 0 \quad (2.6)$$

$$H_h(z) - \nabla h(x)^T dx = 0 \quad (2.7)$$

Uno vé que  $d\mu_\gamma$  puede resolverse explícitamente de la ecuación (2.6) como:

$$d\mu_\gamma = -H_\gamma(z) + \nabla g_\gamma(x)^T dx$$

denotando

$$\Delta(z) = \nabla F(x) + \sum_{i \in \alpha(z)} \mu_i \nabla^2 g_i(x) + \sum_{j=1}^p v_j \nabla^2 h_j(x),$$

$$du_\beta^+ = \max(0, du_\beta),$$

$$q_{\beta}(z) = H_{\beta}(z),$$

$$y = \begin{pmatrix} dx \\ du_{\alpha} \\ dv \end{pmatrix},$$

$$q_y(z) = \begin{pmatrix} H_F(z) \\ H_{\alpha}(z) \\ H_h(z) \end{pmatrix},$$

$$B(z) = \begin{pmatrix} \nabla g_{\beta}(x) \\ 0 \\ 0 \end{pmatrix},$$

$$L(z) = \begin{pmatrix} \Delta(z) & \nabla g_{\alpha}(x) & \nabla h(x) \\ -\nabla g_{\alpha}(x)^T & 0 & 0 \\ -\nabla h(x)^T & 0 & 0 \end{pmatrix},$$

$$P(z) = \begin{pmatrix} L(z) & B(z) \\ -B(z)^T & 0 \end{pmatrix},$$

Entonces de (2.3) – (2.5) y (2.7) tenemos:

$$L(z)y + B(z)du_{\beta}^+ + q_y(z) = 0, \quad (2.8)$$

$$[-B(z)^T y + q_{\beta}(z)]^T du_{\beta}^+ = 0, \quad (2.9)$$

$$-B(z)^T y + q_{\beta}(z) \geq 0, \quad (2.10)$$

$$du_{\beta}^+ \geq 0, \quad (2.11)$$

De las ecuaciones (2.8) – (2.11) definimos un *problema de complementariedad lineal mixto*.

Si  $L(z)$  es una matriz no singular, se puede resolver y en términos de  $du_{\beta}^{+}$ , de (2.8):

$$y = -L^{-1}(z)B(z)du_{\beta}^{+} - L(z)^{-1}q_y(z)$$

Luego substituyendo y en (2.9) – (2.10):

$$\left. \begin{aligned} [B(z)^T L^{-1}(z)B(z)du_{\beta}^{+} + B(z)^T L^{-1}(z)q_y(z) + q_{\beta}(z)]du_{\beta}^{+} &= 0 \\ B(z)^T [L^{-1}(z)B(z)du_{\beta}^{+} + L^{-1}(z)q_y(z)] + q_{\beta}(z) &\geq 0 \\ du_{\beta}^{+} &\geq 0 \end{aligned} \right\} \dots (*)$$

El sistema resultante se convierte entonces en un *problema de complementariedad lineal* en  $du_{\beta}$ .

La matriz en el problema de complementariedad lineal (\*) se convierte entonces en  $B(z)^T L(z)^{-1} B(z)$ , lo cual es el Complemento Schur de  $L(z)$  en  $P(z)$ . Haciendo:

$$M = B(z)^T L^{-1}(z) B(z)$$

$$q = B(z)^T L^{-1}(z)q_y(z) + q_{\beta}(z)$$

$$z = du_{\beta}^{+}$$

$$w = B(z)^T [L^{-1}(z)B(z)du_{\beta}^{+} + L^{-1}(z)q_y(z)] + q_{\beta}(z).$$

Si  $B(z)^T L(z)^{-1} B(z)$  es una  $P$ -Matriz, entonces el problema de complementariedad lineal (\*) tiene una solución única.

Una condición suficiente de (2.8) – (2.11) para tener una solución única es que:

- La matriz  $L(z)$  es no-singular
- El Complemento Schur de  $L(z)$  en  $P(z)$ ,  $B(z)^T L(z)^{-1} B(z)$  es una  $P$ -Matriz.



## 2.1 Regularidad y regularidad débil

### Definición 2.1.1. (Regularidad)

Sea  $z = (x, u, v)^T$  un vector arbitrario en  $\mathbb{R}^{n+m+p}$ . Entonces  $z$  es llamado un vector regular para la función  $H$  definido en (1.14) si:

$$(a) L(z) = \begin{pmatrix} \Delta(z) & \nabla g_{\alpha}(x) & \nabla h(x) \\ -\nabla g_{\alpha}(x)^T & 0 & 0 \\ -\nabla h(x)^T & 0 & 0 \end{pmatrix}$$

es no-singular.

(b) El Complemento Schur de  $L(z)$ , dado por  $B(z)^T L(z)^{-1} B(z)$  es una P-Matriz, donde  $L(z)$  y  $P(z)$  fueron definidos anteriormente.

Luego, si  $z$  es un vector regular entonces el sistema de ecuaciones  $H(z) + H'(z, d) = 0$  tendrá solución única, con lo cual se obtendrá el vector dirección  $d$ .

### Definición 2.1.2.

(Regularidad débil) Sea  $z = (x, u, v)^T$  un vector arbitrario en  $\mathbb{R}^{n+m+p}$ . El vector  $z$  es llamado un vector regular débilmente para la función  $H$  definido en (1.14) si existe una vecindad  $N$  de  $z$  tal que para cada  $z' = (x', u', v')^T \in N$ , la matriz

$$L(z') = \begin{pmatrix} \Delta(z') & \nabla g_{\alpha'}(x') & \nabla h(x') \\ -\nabla g_{\alpha'}(x')^T & 0 & 0 \\ -\nabla h(x')^T & 0 & 0 \end{pmatrix}$$

es no-singular, donde  $\alpha' = \alpha(z')$

Obviamente, si un vector es regular, esto es también regular débilmente. El siguiente ejemplo demuestra sin embargo, que lo inverso no es generalmente cierto.

### Ejemplo 2.1

Consideremos el Problema de Complementariedad No-lineal (NCP) donde

$$F(x) = \begin{pmatrix} x_1^2 + x_1 + 2x_2 - 2 \\ 2x_1^2 - 2x_2 + x_2^2 - 2 \end{pmatrix} \quad (2.12)$$

De acuerdo (1.11) – (1.12), el problema NCP puede ser formulado como el siguiente sistema de ecuaciones:

$$H(z) = \begin{pmatrix} F(x) - u^+ \\ x + u^- \end{pmatrix} = 0$$

donde  $H: \mathbb{R}^{2n} \rightarrow \mathbb{R}^{2n}$ , y  $z = (x, u)^T \in \mathbb{R}^{2n}$ . Luego:

$$H_1(z) = x_1^2 + x_1 + 2x_2 - 2 - u_1^+ = 0$$

$$H_2(z) = 2x_1^2 - 2x_2 + x_2^2 - 2 - u_2^+ = 0$$

$$H_3(z) = x_1 + u_1^- = 0$$

$$H_4(z) = x_2 + u_2^- = 0$$

Sea  $z = (x, u)^T = \left(-\frac{1}{2}, 0, 1, 0\right)^T$  entonces:

$$\alpha(z) = \{1\},$$

$$\beta(z) = \{2\},$$

$$\gamma(z) = \emptyset.$$

En el problema NCP, tenemos:

$$\Delta(z) = \nabla F(x),$$

$$B(z) = \begin{pmatrix} 0 \\ -1 \\ 0 \end{pmatrix},$$

$$L(z) = \begin{pmatrix} \nabla F(x) & -I_{\alpha}^T \\ I_{\alpha} & 0 \end{pmatrix},$$

$$P(z) = \begin{pmatrix} \nabla F(x) & -I_{\alpha}^T & -I_{\beta}^T \\ I_{\alpha} & 0 & 0 \\ I_{\beta} & 0 & 0 \end{pmatrix}$$

donde:  $I_{\alpha}$  de orden  $|\alpha| \times n$  es una submatriz de  $I$ , cuyas filas son indexadas por el conjunto  $\alpha$ .  $I_{\beta}$  definido de manera análoga.

Evaluando, tenemos:

$$\nabla F(x) = \begin{pmatrix} 2x_1 + 1 & 2 \\ 4x_1 & 2x_2 - 2 \end{pmatrix} \Rightarrow \nabla F(x)|_{(-\frac{1}{2}, 0)} = \begin{pmatrix} 0 & 2 \\ -2 & -2 \end{pmatrix},$$

$$I_{\alpha} = (1 \quad 0),$$

$$I_{\beta} = (0 \quad 1),$$

luego:

$$L(z) = \begin{pmatrix} 0 & 2 & -1 \\ -2 & -2 & 0 \\ 1 & 0 & 0 \end{pmatrix} \Rightarrow L^{-1}(z) = \begin{pmatrix} 0 & 0 & 1 \\ 0 & -\frac{1}{2} & -1 \\ -1 & -1 & -2 \end{pmatrix}, \text{ se observa que } L(z) \text{ es no-}$$

singular.

Además:

$$P(z) = \begin{pmatrix} 0 & 2 & -1 & 0 \\ -2 & -2 & 0 & -1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}$$

$$\Rightarrow B(z)^T L(z)^{-1} B(z) = \begin{pmatrix} 0 & -1 & 0 \end{pmatrix} \cdot \begin{pmatrix} 0 & 0 & 1 \\ 0 & -\frac{1}{2} & -1 \\ -1 & -1 & -2 \end{pmatrix} \cdot \begin{pmatrix} 0 \\ -1 \\ 0 \end{pmatrix} = -\frac{1}{2} < 0$$

Entonces el complemento Schur de  $L(z)$  en  $P(z)$ , dado por  $B(z)^T L(z)^{-1} B(z)$  no es una  $P$ -Matriz. Por lo tanto  $z$  no es un vector regular, pues no satisface la condición (b) de la definición de regularidad (2.1.1). Sin embargo  $z$  es un vector regular débilmente, en efecto:

Sólo basta tomar una Vecindad  $N$  de  $z$  tal que  $\forall z' \in N$ ,  $\alpha(z) \subset \alpha(z')$ , pero teniendo en cuenta que:  $|\alpha(z)|=1 \Rightarrow |\alpha(z')|=2$ , debido a que  $u = (u_\alpha, u_\beta, u_\gamma)$ ,  $u \in \mathbb{R}^2$ . Luego  $\beta(z')=\emptyset$ ,  $\gamma(z')=\emptyset$ . Entonces  $L(z')=P'(z')$ , luego por la continuidad de  $P$ ,  $P'(z')$  es no-singular, es decir  $L(z')$  es no singular. Por lo tanto  $z$  es un vector regular débilmente.

De este modo, la regularidad débil es de hecho más débil que la regularidad.

## 2.2 Propiedad descendente

Indicaremos primero que el algoritmo de Newton (amortiguado) descrito en la sección anterior genera una dirección descendente para la función norma  $\theta(z)$ .

### Proposición 2.2.1

Sea la dirección buscada  $d^k$  obtenida al resolver al sistema de ecuaciones:

$$H(z^k) + H'(z^k, d^k) = 0. \quad (2.13)$$

Entonces

$$\theta'(z^k, d^k) = -2\theta(z^k) \leq 0. \quad (2.14)$$

En particular  $\theta'(z^k, d^k) < 0$ , si  $H(z^k) \neq 0$ ,

**El método modificado de Newton (amortiguado) es descrito como sigue:**

1. Si el sistema (2.13) tiene una solución, realizaremos la búsqueda lineal a lo largo de la dirección  $d^k$ ; en otro caso, fijamos  $du_\beta = 0$ , y resolver (2.3), (2.4), (2.7) y (2.6) para la dirección  $dx$ ,  $du_\alpha$ ,  $dv$  y  $du_\gamma$ . Sea  $d^k = (dx, du_\alpha, 0, du_\gamma, dv)$
2. Si  $d^k$  es una dirección descendente de  $\theta$  en  $z$ , realizaremos una búsqueda lineal a lo largo de la dirección  $d^k$  para generar  $z^{k+1} = z^k + \lambda_k d^k$ ; en otro caso, buscamos a lo largo de  $-d^k$  para calcular  $z^{k+1} = z^k - \lambda_k d^k$ .

Cuando (2.13) tiene una solución, una dirección de Newton es obtenida. Esta dirección tiene toda la información de la derivada de  $H$  en  $z$ . Si el sistema no tiene solución, una dirección generada por la modificación anterior no es de Newton y contiene solamente información parcial de la derivada de  $H$  en  $z$ . Sin embargo, si el conjunto  $\beta$  es “pequeño”, lo cual usualmente se produce en la práctica, la dirección generada por el método modificado será una buena aproximación de la dirección de Newton.

El método modificado de Newton (amortiguado) generará una dirección descendente si la regularidad débil y algunas condiciones mild son satisfechas.

### **Proposición 2.2.2**

Sea  $z$  un vector regular débilmente de  $H$  definido por (1.14). Si  $H(z) \neq 0$ , entonces:

Siempre existe una dirección descendente  $d$  en  $z$  para la función norma  $\theta$  a menos que se cumpla las **condiciones mild**:

$$(i) \quad H_\gamma(z) = 0$$

$$(ii) \quad \nabla g_\alpha(x)^T H_F(z) = 0$$

$$(iii) \quad \nabla h(x)^T H_F(z) = 0$$

$$(iv) \quad H_F(z)^T \Delta(z) + g_\alpha(x)^T \nabla g_\alpha(x)^T + g_\beta(x)^T \nabla g_\alpha(x)^T + h(x)^T \nabla h(x)^T = 0$$

$$(v) \quad \nabla g_\beta(x)^T H_F(z) \geq 0$$

$$(vi) \ g_{\beta}(x) \geq 0$$

En particular, el método modificado de Newton (amortiguado) obtendrá una dirección descendente a menos que (i) – (iv) sean satisfechas simultáneamente.

## Conclusión

Una de las dificultades que presenta el método tradicional de Newton es la diferenciabilidad. Cuando la función es consideración no es F-diferenciable en ciertos puntos, el método puede fallar. El algoritmo de Newton (amortiguado) satisface dichas dificultades aplicando el método de Newton a funciones B - diferenciables. La característica se basa fundamentalmente que éste nuevo método identifica aquellas partes en la cual puede causar problemas la no diferenciabilidad de la función (usualmente es el conjunto  $\beta$ ).

Con este trabajo desarrollado estamos en condiciones de realizar experimentos numéricos y comparar los resultados de estos experimentos con los métodos tradicionales como es el método de Newton y se puede mostrar que éste método es más eficiente y robusto que aquellos métodos tradicionales para resolver problemas de complementar.

## Bibliografía

- [1] *M. Avriel. Nonlinear Programming Analysis and Methods. Prentice-Hall. (1976).*
- [2] *M.S. Buzaraa and C.M. Shetty, Nonlinear Programming: Theory and Algorithms. John Wiley and Sons, 1979.*
- [3] *F. Facchinei. "A Simply Constrained optimization reformulation of KKT system arising from variational inequalities". Technical University of Dresden. Institute of Numerical Mathematics. (1997).*
- [4] *M. Fukushima. "Optimization Based Globally Convergent Methods for the Nonlinear Complementary Problem".*
- [5] *M. Fukushima. "Equivalent differentiable optimization problems and descent methods for asymmetric variational inequality problems". Technical report 89007. Department of Applied Mathematics and Physics. Faculty of Engineering. Kyoto University (Kyoto, Japan. 199).*
- [6] *P.T. Harker and J. Pang. "Finite-Dimensional Variational Inequality and*

- Nonlinear Complementarity Problems: A Survey of Theory, Algorithms and Applications*". *Mathematical Programming B*-48. (1990) 1-60.
- [7] **P.T. Harker and B. Xiao**. "Newton's method for the nonlinear complementary problem: A B-differentiable equation approach". *Mathematical programming* 48 (1990) 339-357.
- [8] **D. Kinderlehrer**. "An Introduction to variational Inequalities and their applications". Academic Press. (1980).
- [9] **Malcolm Keswell**. "Notes on Second Order "Sufficient" Conditions for Characterising Extrema of Bivariate and Multivariate Functions".
- [10] **J.S. Pang**. "Minimization of Locally Lipschitzian Functions". *Society for Industrial and Applied Mathematics* (1991) 57-82.
- [11] **J.S. Pang**. "Newton's method for B-differentiable equations". *Mathematics of Operations Research* 15 (1990) 311-341.
- [12] **J.S. Pang**. "A B-differentiable equation based, globally, and locally quadratically convergent algorithm for nonlinear programs, complementarity and variational inequality problems". *Mathematical Programming* 51 (1991) 101-131.
- [13] **S.M. Robinson**. "False Numerical Convergence in some Generalized Newton Methods". Department of Industrial Engineering, University of Wisconsin (Madison, WI, 1995).
- [14] **S.M. Robinson**. "An implicit-function theorem for a class of nonsmooth functions". Manuscript. Department of Industrial Engineering, 1991.
- [15] **A. Shapiro**. "On concepts of directional differentiability". Research Report 73/88(18). Department of Mathematics and Applied Mathematics, University of South Africa (Pretoria, South Africa. 1988).
- [16] **B. Xiao and P. Harker**. "A nonsmooth Newton method for variational inequalities, I: Theory". Department of System Engineering, University of Pennsylvania, Philadelphia. 1993.

# Una prueba general de la buena definición del método lagrangeano aumentado

---

*Yna Consuelo Rezza Espinoza\**

## **RESUMEN**

En este trabajo presentamos una demostración de la **buena definición** del Método Lagrangeano Aumentado con penalidades  $P_i \in \mathcal{P}$  la cual incluye casi todas las pruebas existentes en la literatura. Las penalidades que describimos consideran dos subfamilias; una, continuamente diferenciable y estrictamente convexa (entre otras características) y la otra, continuamente diferenciable, estrictamente convexa en  $[-b, +\infty)$ , para algún  $b > 0$  y constante en  $(-\infty, -b]$ .

Las pruebas de buena definición realizadas por Rockafellar [10], Bertsekas [1], Polyak y Teboulle [7] y Gonzaga y Castillo [2] vendrían a ser casos particulares.

## **Introducción**

Problemas no lineales de optimización surgen espontáneamente en muchos campos de aplicación. Uno de los más estudiados es el siguiente:

---

\* Universidad Federal de Santa Catarina  
Universidad Nacional de Ingeniería



minimizar  $f(x)$

s.a.  $g_i(x) \leq 0$ ,  $i = 1, \dots, m$

$x \in \mathbb{R}^n$

donde  $f: \mathbb{R}^n \rightarrow \mathbb{R} \cup \{+\infty\}$ ,  $g_i: \mathbb{R}^n \rightarrow \mathbb{R} \cup \{+\infty\}$  para  $i = 1, \dots, m$  son funciones convexas y cerradas.

Para resolver este problema de optimización convexa apareció a fines de la década de los 70 un método que se ha venido aplicando con éxito hasta el día de hoy; estamos refiriéndonos al Método Lagrangeano Aumentado. Basado, como muchos otros métodos, en la sustitución del problema inicial por una secuencia de subproblemas, el primer objetivo a alcanzar será garantizar, la **buena definición** de la sucesión formada por alguna solución de cada subproblema.

Rockafellar [10], demostró esto mediante el uso de penalidades Estándar, Bertsekas [1] la garantizó para penalidades del tipo Exponencial. Otras pruebas de la buena definición del Método Lagrangeano Aumentado se encuentran en [7] y [2].

La prueba de buena definición que presentamos, incluye todos estos casos. Nosotros definimos las penalidades  $P_i \in \mathcal{P}$  considerando dos subfamilias, una continuamente diferenciable y estrictamente convexa y la otra, continuamente diferenciable, estrictamente convexa en  $[-b, +\infty)$  para algún  $b > 0$  y constante en  $(-\infty, -b]$ .

## Notación

$\text{int } S$  Interior del conjunto  $S$ .

$\text{ri } S$  Interior relativo del conjunto  $S$ .

$\mathbb{R}_+$  Conjunto de números reales no negativos.

$\mathbb{R}_{++}$  Conjunto de números reales negativos.

$\mathbb{R}_+^m$  Conjunto de puntos de  $\mathbb{R}^m$  cuyas componentes son no negativas.

$\mathbb{R}_{++}^m$  Conjunto de puntos de  $\mathbb{R}^m$  cuyas componentes son negativas.

$\partial f(x)$  Conjunto de subgradientes de la función de  $f$  en el punto  $x$ .

## La buena definición del Método Lagrangeano aumentado

### Definición 2.1

Se dice que la función  $f : \mathbb{R}^n \rightarrow \mathbb{R} \cup \{+\infty\}$  no idénticamente  $+\infty$  es convexa si para todo  $(x, x') \in \mathbb{R}^n \times \mathbb{R}^n$  y todo  $\alpha \in (0, 1)$  tenemos:

$$f(\alpha x + (1-\alpha)x') \leq \alpha f(x) + (1-\alpha)f(x')$$

considerando esta desigualdad en  $f \in \mathbb{R} \cup \{+\infty\}$ .

Esta clase de funciones será denotada así  $f \in \text{Conv } \mathbb{R}^n$ .

### Definición 2.2

La función  $f : \mathbb{R}^n \rightarrow \mathbb{R} \cup \{+\infty\}$  se dice cerrada, si para cada  $x \in \mathbb{R}^n$  se verifica

$$\liminf_{y \rightarrow x} f(y) \geq f(x).$$

Si  $f \in \text{Conv } \mathbb{R}^n$  y además  $f$  es cerrada, denotaremos  $f \in \overline{\text{Conv } \mathbb{R}^n}$ .

Consideremos el siguiente problema de optimización convexa:

$$\begin{aligned} &\text{minimizar } f(x) \\ &\text{s.a. } g_i(x) \leq 0, \quad i = 1, \dots, m \\ &\quad x \in \mathbb{R}^n \end{aligned}$$

donde  $f : \mathbb{R}^n \rightarrow \mathbb{R} \cup \{+\infty\}$ ,  $g_i : \mathbb{R}^n \rightarrow \mathbb{R} \cup \{+\infty\}$  para  $i = 1, \dots, m$  son funciones convexas y cerradas.

### Hipótesis 2.3

Supongamos que el conjunto de soluciones óptimas del problema  $(\hat{P})$  es no vacío y limitado.

**Hipótesis 2.4 (Condición de Calificación de Slater).**

Supongamos que existe  $\bar{x} \in \mathbb{R}^n$  tal que  $g_i(\bar{x}) < 0 \quad \forall i=1, \dots, m$ .

**Definición 2.5**

Dado  $b > 0$  (posiblemente  $+\infty$ ), denotemos por  $\mathcal{P}$  a la familia de funciones  $P: \mathbb{R} \times \mathbb{R}_{++} \rightarrow \mathbb{R}$  o  $P: \mathbb{R} \times \mathbb{R}_+ \rightarrow \mathbb{R}$  que satisfacen lo siguiente:  
Si  $b = +\infty$  (respectivamente  $0 < b < +\infty$ )

1.-  $P(., u)$  es continuamente diferenciable en  $\mathbb{R}$ , estrictamente convexa en  $[-b, +\infty)$  y constante en  $(-\infty, -b]$ .

2.-  $P(0, u) = 0$ ,  $\frac{\partial P}{\partial t}(0, u) = u$

3.-  $\lim_{t \rightarrow -\infty} \frac{\partial P}{\partial t}(t, u) = 0$  y

4.-  $\lim_{t \rightarrow +\infty} \frac{\partial P}{\partial t}(t, u) = +\infty$

para cualquier  $u > 0$  y  $\forall t \in \mathbb{R}$  (respectivamente para cualquier  $u \geq 0$  y  $\forall t \in \mathbb{R}$ ).

Por comodidad de aquí en adelante escribiremos  $P'(., u)$  en lugar de  $\frac{\partial P}{\partial t}(., u)$ , para cualquier  $u \geq 0$ .

**Las funciones Lagrangeano y Lagrangeano Aumentado**

La función Lagrangeano asociada al problema  $(\hat{P})$  es definida por

$$x \in \mathbb{R}^n, \mu \in \mathbb{R}^m \rightarrow l(x, \mu) = f(x) + \sum_{i=1}^m \mu_i g_i(x)$$

aparece en la teoría de la optimización al tratar de las condiciones necesarias y

suficientes que una solución del problema  $(\hat{P})$  verifica. Como veremos más adelante, estas condiciones (y la función lagrangeano clásico) serán parte fundamental en la construcción del método Lagrangeano Aumentado en estudio. De manera similar definimos, asociada al problema  $(\hat{P})$ ,  $L$ , la función Lagrangeano aumentado (con penalidades  $P_i \in \mathcal{P}$ ) dada por:

$$x \in \mathbb{R}^n, \mu \in \mathbb{R}^m, \lambda > 0 \rightarrow L(x, \mu, \lambda) = f(x) + \lambda \sum_{i=1}^m P_i \left( \frac{g_i(x)}{\lambda}, \mu_i \right) \quad (2.1)$$

donde las  $m$  funciones  $P_i: \mathbb{R} \times \mathbb{R}_{++} \rightarrow \mathbb{R}$  (o las  $m$  funciones  $P_i: \mathbb{R} \times \mathbb{R}_+ \rightarrow \mathbb{R}$ ) llamadas **penalidades** pertenecerán a la familiar  $\mathcal{P}$  definida antes.

La primera función Lagrangeano aumentado introducida tenía aplicación a problemas con restricciones de igualdad (ver [5] y [8]). Rockafellar [10] generalizó la metodología para el caso donde las restricciones son de desigualdad. Esta primera función L.A. es dada por:

$$x \in \mathbb{R}^n, \mu \in \mathbb{R}^m, \lambda > 0 \rightarrow L(x, \mu, \lambda) = f(x) + \lambda \sum_{i=1}^m \frac{1}{2} \left[ \left( \max \left\{ 0, \mu_i, \frac{g_i(x)}{\lambda} \right\} \right)^2 - \mu_i^2 \right].$$

## Resultados preliminares a la descripción del método

### Definición 2.6

Sea  $C \subset \mathbb{R}^n$  un convexo. La aplicación  $F: C \rightarrow \mathbb{R}^n$  se dice monótona (respectivamente estrictamente monótona) en  $C$  si  $\forall x, x' \in C$ ,

$$(F(x) - F(x'))^t (x - x') \geq 0$$

(respectivamente  $(F(x) - F(x'))^t (x - x') > 0$ , cuando  $x \neq x'$ ).

### Teorema 2.7

Sea  $f$  una función diferenciable en un abierto  $\Omega \subset \mathbb{R}^n$  y sea  $C \subset \Omega$  un convexo.

Luego  $f$  es convexa (respectivamente estrictamente convexa) en  $C$  si y sólo si  $\nabla f$  es monótono (respectivamente estrictamente monótono) en  $C$ .

### Prueba

Ver Teorema 4.1.4 pag 185 en [3].

### Observación 2.8

Se  $f: \mathbb{R} \rightarrow \mathbb{R}$  convexa (respectivamente estrictamente convexa) y diferenciable. Para  $\mu \geq 0$  (respectivamente  $\mu > 0$ ), si  $f'(0) = \mu$  y  $\lim_{t \rightarrow -\infty} f'(t) = 0$  entonces

$$f'(t) \geq 0 \text{ (respectivamente } f'(t) > 0 \text{) para todo } t \in \mathbb{R}.$$

### Prueba

a) Para  $f$  convexa. Dado que  $f'(t) = \mu \geq 0$  para  $t=0$ , analizaremos dos casos.

i) Si  $t \geq 0$  entonces por el Teorema 2.7  $f'(t) \geq f'(0)$ . Así  $f'(t) \geq \mu \geq 0$ .

ii) Caso  $t < 0$ . Supongamos que existe  $\hat{t} < 0$  tal que  $f'(\hat{t}) < 0$  entonces, para todo  $t < \hat{t}$  tenemos  $f'(t) \leq f'(\hat{t}) < 0$  por el mismo teorema. Haciendo tender  $t \rightarrow -\infty$  obtenemos

$$\lim_{t \rightarrow -\infty} f'(t) \leq f'(\hat{t}) < 0$$

lo que contradice la hipótesis. Luego, para todo  $t < 0$ ,  $f'(t) \geq 0$ . Por tanto,

$$f'(x) \geq 0 \quad \forall t \in \mathbb{R}.$$

b) Para  $f$  estrictamente convexa. Supongamos que existe  $\hat{t} \in \mathbb{R}$  tal que  $f'(\hat{t}) = 0$ . Entonces, para todo  $t < \hat{t}$ , nuevamente por el Teorema 2.7, tenemos  $f'(t) < f'(\hat{t})$ . Así  $f'(t) < f'(\hat{t}) = 0$ .

Por el resultado anterior (caso  $f$  convexa) sabemos que  $f'(t) \geq 0 \quad \forall t \in \mathbb{R}$ . Luego,  $0 \leq f'(t) < 0$ , lo que es una contradicción. Por tanto, no existe  $t \in \mathbb{R}$  tal que  $f'(t) = 0$ , o sea, retiramos la igualdad en la conclusión de la Observación.

### Proposición 2.9

Sea  $f \in \text{Conv } \mathbb{R}^n$  (respectivamente  $\overline{\text{Conv}} \mathbb{R}^n$ ) y sea  $g \in \text{Conv } \mathbb{R}^n$  (respectivamente  $\overline{\text{Conv}} \mathbb{R}^n$ ) creciente. Supongamos que existe  $x^0 \in \mathbb{R}^n$  tal que  $f(x^0) \in \text{dom } g$ . Luego, la función compuesta  $x \mapsto g(f(x))$  está en  $\text{Conv } \mathbb{R}^n$  (respectivamente  $\overline{\text{Conv}} \mathbb{R}^n$ ).

### Prueba

Ver Proposición 2.1.8, pag 160, en [3].

### Observación 2.10

Para cada  $u \geq 0$  y  $\lambda > 0$  fijos. Considere  $P \in \mathcal{P}$  y  $g: \mathbb{R}^n \rightarrow \mathbb{R} \cup \{+\infty\}$  una función convexa y cerrada. Entonces:

$$i) \quad \frac{\partial P}{\partial t}(t, u) = P'_u(t) \geq 0 \quad \forall t \in \mathbb{R}. \text{ O sea, } P_u = P(\cdot, u) \text{ es creciente.}$$

$$ii) \quad P(g(\cdot)/\lambda, u) = P_u \circ g/\lambda \text{ es convexa y cerrada.}$$

Además, si  $P \in \mathcal{P}$  es tal que para cualquier  $u > 0$ ,  $P(\cdot, u)$  es estrictamente convexa (el caso en que  $b = +\infty$ ), entonces  $P'_u > 0$ ,  $\forall t \in \mathbb{R}$ . O sea,  $P_u$  será estrictamente creciente.

### Prueba

i) Como  $P \in \mathcal{P}$ , para  $u \geq 0$  fijo  $P(\cdot, u) = P_u(\cdot)$  es convexa y diferenciable en  $\mathbb{R}$ . Además de eso  $P_u$  verifica las condiciones de la Observación 2.8 entonces  $P'_u \geq 0 \quad \forall t \in \mathbb{R}$ . Luego,  $P_u$  será creciente.

ii) Dado que  $g$  es convexa y cerrada, para  $\lambda > 0$  fijo,  $g/\lambda$  también lo será. Como para  $u \geq 0$  fijo,  $P_u$  es convexa e creciente, tenemos por la Proposición 2.9 que la compuesta  $P_u \circ g/\lambda$  también estará en  $\overline{\text{Conv}} \mathbb{R}^n$ .

La última afirmación se obtiene del caso alternativo en la Observación 2.8.

**Observación 2.11**

Para  $\mu \geq 0$  en  $\mathbb{R}^m$  y  $\lambda > 0$  fijos, consideremos la función Lagrangeano aumentado con penalidades  $P_i \in \mathcal{P}$  definida en (2.1). Entonces,

$$x \in \mathbb{R}^n \mapsto L(x, \mu, \lambda) = L_{\mu, \lambda}(x) = f(x) + \lambda \sum_{i=1}^m P_i \left( \frac{g_i(x)}{\lambda}, \mu_i \right)$$

en convexa y cerrada.

**Prueba**

Para  $\mu \geq 0$  en  $\mathbb{R}^m$  fijo, tenemos  $\mu_i \geq 0 \quad \forall i=1, \dots, m$ . Entonces, como cada  $P_i \in \mathcal{P}$  y  $g_i$  son convexas y cerradas, por la Proposición 2.9,  $P_{\mu_i} \circ g_i / \lambda$  estará en  $\overline{\text{Conv}} \mathbb{R}^n$ , para  $\lambda > 0$  fijo.

Finalmente, como  $L_{\mu, \lambda}$  es una combinación positiva de funciones convexas y cerradas,  $L_{\mu, \lambda}$  también estará en  $\overline{\text{Conv}} \mathbb{R}^n$ .

**Proposición 2.12**

Para  $f \in \text{Conv} \mathbb{R}^n$ , se tiene:

$$x \in \mathbb{R}^n \text{ minimiza } f \text{ si y solo si } 0 \in \partial f(x).$$

**Prueba**

Consecuencia del Corolario 1.4.4, pag. 48 dado en [4].

**Corolario 2.13**

Sean  $f_1, f_2$ , dos funciones convexas y cerradas. Asumamos que  $\text{ridom} f_1 \cap \text{ridom} f_2 \neq \emptyset$ . Luego

$$\partial(f_1 + f_2)(x) = \partial f_1(x) + \partial f_2(x) \text{ para todo } x \in \text{dom} f_1 \cap \text{dom} f_2.$$

### Prueba

Ver Corolario 3.1.2, pag. 114 en [4].

### Teorema 2.14

Sean  $f : \mathbb{R}^n \rightarrow \mathbb{R} \cup \{+\infty\}$  convexa y  $g \in \overline{\text{Conv}} \mathbb{R}^n$  creciente. Asumamos que  $f(\mathbb{R}^n) \cap \text{int dom} g \neq \emptyset$ . Luego, para todo  $x$  tal que  $f(x) \in \text{dom} g$

$s \in \partial(g \circ f)(x)$  si y solo si  $\exists \alpha \geq 0$  tal que  $s \in \partial(\alpha f)(x), \alpha \in \partial g(f(x))$ .

Si además,  $g$  es diferenciable en  $f(x)$ , la relación anterior puede expresarse por:

$$\partial(g \circ f)(x) = \partial g(f(x)) \partial f(x).$$

### Prueba

Consecuencia de los Teoremas 3.6.1, pags 125 y 126, en [4] y 4.3.1., pag 265 en [3].

### Definición 2.15

Un punto  $\bar{x} \in \mathbb{R}^n$  es llamado un punto de Karush Kuhn Tucker (K.K.T) si existe algún punto  $\bar{u} \in \mathbb{R}^m$  tal que se verifica

$$a) \quad 0 \in \partial f(x) + \sum_{i=1}^m \bar{u}_i \partial g_i(\bar{x}) = \partial l(\bar{x}, \bar{u})$$

$$b) \quad \bar{u} \geq 0$$

$$c) \quad g_i(\bar{x}) \leq 0 \quad \forall \quad i=1, \dots, m$$

$$d) \quad \bar{u}_i g_i(\bar{x}) = 0 \quad \forall \quad i=1, \dots, m$$

Si se verifica apenas a) diremos que  $(\bar{x}, \bar{u})$  cumplen la condición de Lagrange.



**Definición 2.16**

Una función definida de la siguiente manera:  $x \in \mathbb{R}^n \mapsto \langle s, x \rangle + k$ , para algún  $s \in \mathbb{R}^n$  y algún  $k \in \mathbb{R}$ , es llamada función afin.

Sea  $X = \{ x \in \mathbb{R}^n : g_i(x) \leq 0 \}$  el conjunto de puntos viables para el problema  $(\hat{P})$ . Denotemos por  $J_a$  al conjunto de índices correspondientes a las restricciones  $g_i$  afines:

$$J_a = \{ i=1, \dots, m : g_i \text{ es una función afin} \}$$

**Definición 2.17**

Decimos que  $X$  satisface la hipótesis flaca de Slater si existe un punto  $x^0 \in \mathbb{R}^n$  en el cual:

i)  $g_i(x^0) \leq 0$  para  $i \in J_a$

ii)  $g_i(x^0) < 0$  para  $i \notin J_a$

Al asumir la Hipótesis 2.4, se estará satisfaciendo la hipótesis flaca de Slater.

**Teorema 2.18**

Supongamos que se verifica la hipótesis flaca de Slater. Entonces, para cualquier  $\bar{x} \in \mathbb{R}^n$ , el hecho de que  $\bar{x}$  sea un punto de Karush Kuhn Tucker es suficiente y necesario para que  $\bar{x}$  sea una solución del problema  $(\hat{P})$ .

**Prueba**

Ver Teorema 2.2.5, pag 310, en [3]

## Descripción del Método Lagrangeano Aumentado

Consideremos la función Lagrangeano Aumentado con penalidades  $P_i \in \mathcal{P}$  asociada al problema  $(\hat{P})$  y  $\{\lambda_k\} \subset \mathbb{R}$  una secuencia limitada de números reales positivos.

El Método Lagrangeano Aumentado para la optimización convexa intenta aproximar un punto de K.K.T. (que bajo la Hipótesis 2.3 y por el Teorema 2.18, será también una solución del problema  $(\hat{P})$ ) mediante la secuencia  $\{x^k\} \subset \mathbb{R}^n$  generada al resolver iterativamente, partiendo de algún punto  $\mu^0 \in \mathbb{R}_+^m$  ( $\mu^0 \in \mathbb{R}_{++}^m$ ) dado, problemas del tipo:

$$\underset{x \in \mathbb{R}^n}{\text{minimizar}} L(x, \mu_k, \lambda_k) = f(x) + \lambda_k \sum_{i=1}^m P_i \left( \frac{g_i(x)}{\lambda_k}, \mu_i^k \right) \quad (\hat{P})$$

donde la actualización de  $\mu^k$  se realiza por la fórmula:

$$\mu_i^{k+1} = \frac{\partial P_i}{\partial y_i}(y_i^{k+1}, \mu_i^k), \quad y_i^{k+1} = \frac{g_i(x^{k+1})}{\lambda_k}, \quad \forall i = 1, \dots, m \quad (2.3)$$

( $x^{k+1}$  denota una solución del problema  $(\hat{P}_k)$ ).

La idea es procurar verificar las condiciones dadas por la Definición 2.15. Primero imponemos la verificación de la condición de Lagrange para todos los iterados. Exploremos esto.

Para  $x^{k+1}$ , una solución del problema  $(\hat{P}_k)$  (que por ahora supondremos que existe), tenemos por la Observación 2.11 y la Proposición 2.12

$$0 \in \partial_{\mu^k, \lambda_k} L(x^{k+1}) \quad (2.4)$$

Evaluando la relación (2.1) en  $(\mu_k, \lambda_k)$  obtenemos:

$$L_{\mu^k, \lambda_k}(x) = f(x) + \lambda_k \sum_{i=1}^m P_{i_{\mu_i^k}} \left( \frac{g_i(x)}{\lambda_k} \right),$$

usando ahora el Corolario 2.13:

$$\partial L_{\mu^k, \lambda_k}(x) = \partial f(x) + \lambda_k \sum_{i=1}^m \partial \left[ P_{i_{\mu_i^k}} \left( \frac{g_i(x)}{\lambda_k} \right) \right]. \quad (2.5)$$

Como  $P_{i_{\mu_i^k}}$ , es creciente (Observación 2.10 i) y diferenciable en  $\mathbb{R}$ , por el Teorema 2.14,

$$\partial \left( P_{i_{\mu_i^k}} \circ \frac{g_i(x)}{\lambda_k} \right)(x) = \frac{\partial}{\partial y_i} P_{i_{\mu_i^k}} \left( \frac{g_i(x)}{\lambda_k} \right) \frac{\partial g_i(x)}{\lambda_k}, \quad y_i = \frac{g_i(x)}{\lambda_k}$$

Sustituyendo esto en (2.5), evaluando en  $x^{k+1}$  y retomando (2.4)

$$0 \in \partial L_{\mu^k, \lambda_k}(x^{k+1}) = \partial f(x^{k+1}) + \sum_{i=1}^m \underbrace{\frac{\partial}{\partial y_i} P_{i_{\mu_i^k}} \left( \frac{g_i(x^{k+1})}{\lambda_k} \right)}_{\mu_i^{k+1}} \partial g_i(x^{k+1}) \quad (2.6)$$

Así  $(x^{k+1}, \mu^{k+1})$  verifican la condición de Lagrange, lo que garantiza que un punto límite de la secuencia  $\{x^k, u^k\}$  (si existe) también satisfecerá dicha condición.

Por otro lado, usando la Observación 2.10 i (con  $\mu = \mu^k \geq 0$ ,  $\lambda = \lambda_k > 0$  y  $\bar{x} = x^{k+1}$ , para un  $k \in \mathbb{N}$  dado) y la fórmula (2.3) obtenemos  $\mu^{k+1} \geq 0$  (respectivamente  $\mu^{k+1} > 0$  para todo  $k \in \mathbb{N}$ , para penalidades  $P_i(\cdot, \mu_i^k)$  estrictamente convexas).

Luego, la viabilidad dual se verifica para un punto límite  $\bar{u}$  de la secuencia  $\{\mu^k\}$ .

Viabilidad primal y complementaridad (ítems c e d de la Definición 2.15 respectivamente), no se satisfacen a lo largo de los iterados. No entrando deberán verificarse en un punto límite  $(\bar{u}, \bar{x})$  (esto sucede por ejemplo en el caso de las penalidades de Polyak y Teboulle [7]).

Para conseguir estos objetivos, así como para garantizar la existencia de soluciones del problema  $(P_k)$  y la convergencia de las secuencias  $\{x^k\}$  y  $\{\mu^k\}$ , deberemos asumir algunas hipótesis, pero ese no es nuestro objetivo. Nosotros estamos interesados específicamente en probar la buena definición de la secuencia  $\{x^k\} \subset \mathbb{R}^n$  generada por el Método Lagrangeano Aumentado con penalidades  $P_i \in \mathcal{P}$ .

Esta prueba incluye casi todos los casos existentes en la literatura, tales como los que se encuentran en [10] (donde se usa la penalidad Estándar), en [1] (que usa la penalidad Exponencial), en [7] (con penalidades Estrictamente Convexas) y en [2] que considera penalidades con Cambio de Variable.

Comenzaremos presentando algunas definiciones y resultados que serán de utilidad.

### Definiciones y Resultados Preliminares

Para  $f : \mathbb{R}^n \rightarrow \mathbb{R} \cup \{+\infty\}$  y algún  $r \in \mathbb{R}^n$  denotemos por  $S_r(f)$  a los conjuntos de nivel de  $f$  dado por:  $S_r(f) = \{x \in \mathbb{R}^n : f(x) \leq r\}$ .

Sea  $h : \mathbb{R}^n \rightarrow \mathbb{R} \cup \{+\infty\}$  una función convexa cerrada, recordemos su función de recesión  $h'_\infty(d)$

$$d \in \mathbb{R}^n \mapsto h'_\infty(d) = \lim_{\lambda \rightarrow \infty} \frac{h(x + \lambda d) - h(x)}{\lambda}, \text{ para } x \in \text{dom } h.$$

### Proposición 2.19

Para  $f \in \overline{\text{Conv}} \mathbb{R}^n$ , los siguientes items son equivalentes:

- i) existe  $r \in \mathbb{R}$  para el  $S_r(f)$  es no vacío y compacto,
- ii) todos los conjuntos de nivel del  $f$  son compactos,
- iii)  $f'_\infty(d) > 0$  para todo  $d \neq 0$  en  $\mathbb{R}^n$ .

### Prueba

Ver Proposición 3.2.5, pag. 180 en [3].

### Definición 2.20

Una dirección  $d \neq 0$  en  $\mathbb{R}^n$  es una dirección de recesión de  $h$  si  $h'_\infty(d) \leq 0$ , donde  $h'_\infty(\cdot)$ , denota la función de recesión de  $h$ .

### Observación 2.21

Si el conjunto de soluciones óptimas de problemas  $(\hat{P})$  es no vacío y limitado (Hipótesis 2.3), entonces las funciones  $f, g_1, \dots, g_m$  no tienen ninguna dirección de recesión común.

### Prueba

Por contradicción. Supongamos que  $h$  es una dirección de recesión común. Entonces, para  $x^*$ , una solución óptima, tenemos que  $\forall \lambda \geq 0$ ,  $x^* + \lambda h$  será viable y  $f(x^* + \lambda h) \leq f(x^*)$ , lo que es una contradicción, pues el conjunto óptimo es limitado. Luego, no existe dirección de recesión común.

### Proposición 2.22

Sean  $f_1, \dots, f_m$   $m$  funciones en  $\overline{\text{Conv}} \mathbb{R}^n$  y  $t_1, \dots, t_m$   $m$  números positivos. Asumamos que existe  $x^0 \in \mathbb{R}^n$  en el  $f_j$  es finito  $\forall j=1, \dots, m$ . Luego, para  $f = \sum_{j=1}^m t_j f_j$  tenemos  $f'_\infty = \sum_{j=1}^m t_j f''_{j_\infty}$ .

### Prueba

Ver Proposición 3.2.9, pag 182 en [3].

### Observación 2.23

Sea  $f : \mathbb{R}^n \rightarrow \mathbb{R} \cup \{+\infty\}$  una función convexa y cerrada. Luego para  $\bar{x} \in \mathbb{R}^n$  y  $\forall h \in \mathbb{R}^n$  tal que  $h \neq 0$ ,

$$f'_\infty(h) = \lim_{\lambda \rightarrow \infty} f'(\bar{x} + \lambda h, h)$$

donde  $f'(x, h)$  denota la derivada direccional de  $f$  en  $x$ , en la dirección  $h$ .

### Prueba

Para todo  $x \in \mathbb{R}^n$  y para cualquier  $h \in \mathbb{R}^n$  tal que  $h \neq 0$ , por la definición de  $f'(x, h)$  tenemos  $f(x + \lambda h) \geq f(x) + \lambda f'(x, h)$  para todo  $\forall \lambda > 0$ . Entonces

$$\lim_{\lambda \rightarrow \infty} \frac{f(x + \lambda h) - f(x)}{\lambda} \geq f'(x, h)$$

O sea,  $f'_\infty(h) \geq f'(x, h)$ . Haciendo  $x = \bar{x} + \lambda h$  para  $\bar{x} \in \mathbb{R}^n$  fijo y para todo  $\lambda > 0$  tenemos:

$$f'_\infty(h) \geq f'(\bar{x} + \lambda h, h) \quad \forall \lambda > 0$$

Tomando límite cuando  $\lambda \rightarrow \infty$

$$f'_\infty(h) \geq \lim_{\lambda \rightarrow \infty} f'(\bar{x} + \lambda h, h) \quad \forall \lambda > 0$$

Para  $y = \bar{x} + \lambda h \in \mathbb{R}^n \quad \forall \lambda > 0$  entonces  $f(\bar{x}) \geq f(y) + \lambda f'(y, -h)$ . Sustituyendo  $y$ ,  $f(\bar{x}) \geq f(\bar{x} + \lambda h) + \lambda f'(y, -h)$ . Pero, como  $f$  es convexa se verifica también  $f'(y, h) \geq -f'(y, -h)$ .

Sustituyendo esto en la relación anterior:

$$f(\bar{x} + \lambda h) - f(\bar{x}) \leq \lambda f'(\bar{x} + \lambda h, h) \quad \forall \lambda > 0.$$

Tomando límite cuando  $\lambda \rightarrow \infty$ :

$$\lim_{\lambda \rightarrow \infty} \frac{f(\bar{x} + \lambda h) - f(x)}{\lambda} \leq \lim_{\lambda \rightarrow \infty} f'(\bar{x} + \lambda h, h)$$

o sea,

$$f'_\infty(h) \leq \lim_{\lambda \rightarrow \infty} f'(\bar{x} + \lambda h, h)$$

Luego tenemos probada la igualdad.

### Observación 2.24

Sean  $f : \mathbb{R}^n \rightarrow \mathbb{R} \cup \{+\infty\}$  una función convexa y cerrada,  $h \in \mathbb{R}^n$  y  $x \in \mathbb{R}^n$ .  
Entonces

$$f'(x, h) = \max_{s \in \partial f(x)} \{\langle s, h \rangle\}$$

donde  $\partial f(x)$  es no vacío compacto y convexo.

### Prueba

Consecuencia del resultado (2.0.1), pag 102, en [4].

### Observación 2.25

Para  $P : \mathbb{R} \rightarrow \mathbb{R}$  una función convexa diferenciable y  $g : \mathbb{R}^n \rightarrow \mathbb{R} \cup \{+\infty\}$  una función convexa cerrada,

definamos la compuesta  $x \in \mathbb{R}^n \mapsto f(x) = P(g(x))$ .

Si  $P'(t) \geq 0 \quad \forall t \in \mathbb{R}$  entonces

$$f'(x, h) = P'(g(x))g'(x, h) \quad \forall h \in \mathbb{R}^n, \quad x \in \mathbb{R}^n.$$

## Prueba

Para  $x \in \mathbb{R}^n$ , arbitrario,  $h \in \mathbb{R}^n$  fijo, de la Observación 2.24 tenemos:

$$\begin{aligned} f'(x, h) &= \max \{ \langle s, h \rangle, s \in \partial f(x) \} \\ &= \max \{ s^t h : s = P'(g(x))\gamma, \gamma \in \partial g(x) \} \\ &= \max \{ P'(g(x))\gamma^t h : \gamma \in \partial g(x) \} \\ &= P'(g(x)) \max \{ \gamma^t h : \gamma \in \partial g(x) \} \end{aligned}$$

usando nuevamente la Observación 2.24 ahora con  $g: f(x, h) = P'(g(x))g'(x, h)$ , que es lo que queríamos probar.

## La buena definición del Algoritmo Lagrangeano Aumentado con penalidades $P_i \in \mathcal{P}$

Consideremos el problema  $(p)$  bajo la Hipótesis 2.3. Sea  $\{x^k\} \subset \mathbb{R}^n$  una secuencia generada por el algoritmo Lagrangeano Aumentado con penalidades  $P_i \in \mathcal{P}$  siguiente:

$$\mu^0 \in \mathbb{R}_+^m \left( \mu^0 \in \mathbb{R}_{++}^m \right) \quad (2.7)$$

$$x^{k+1} = \arg \min_{x \in \mathbb{R}^n} \{ L(x, \mu^k, \lambda_k) \} \quad (2.8)$$

$$\mu_i^{k+1} = \frac{\partial P_i}{\partial y_i}(y_i^{k+1}, \mu_i^{k+1}) \quad \forall i = 1, \dots, m \quad \forall k \in \mathbb{N} \quad (2.9)$$

donde  $\{\lambda_k\}$  es una secuencia limitada de números reales positivos,  $L$  es la función Lagrangeano Aumentado con penalidades  $P_i \in \mathcal{P}$  y  $y_i^{k+1} = g_i(x^{k+1})/\lambda_k$  para  $i=1, \dots, m$ .

## Afirmación 2.26

La secuencia  $\{x^k\}$  está bien definida.



## Prueba

Dados  $\lambda_k > 0$  y  $\mu^k \geq 0$  ( $\mu^k > 0$ ),  $k \in \mathbb{N}$  arbitrario, probaremos que  $L_{\mu^k, \lambda_k}$  no tiene dirección de recesión (lo que según la Proposición 2.19 será equivalente a afirmar que los subconjuntos de nivel de  $L_{\mu^k, \lambda_k}$  son compactos. Luego  $L_{\mu^k, \lambda_k}$  alcanza su mínimo en  $\mathbb{R}^n$ ).

De la definición de  $L_{\mu^k, \lambda_k}$ , como  $f$  y  $P_{\mu^k} \circ g_i / \lambda_k$  están en  $\overline{\text{Conv}} \mathbb{R}^n$ , por la Proposición 2.22 tenemos:

$$L'_{\infty}(z, \mu^k, \lambda_k) = f'_{\infty}(z) + \lambda_k \sum_{i=1}^m h'_{i_{\infty}}(z) \quad \forall z \in \mathbb{R}^n$$

donde  $h_i(x) = P_i(g_i(x) / \lambda_k, \mu_i^k) \forall i = 1, \dots, m$  o también usando la Observación 2.23:

$$\lim_{\lambda \rightarrow \infty} L'_{\mu^k, \lambda_k}(x + \lambda z, z) = \lim_{\lambda \rightarrow \infty} f'(x + \lambda z, z) + \lambda_k \sum_{i=1}^m \lim_{\lambda \rightarrow \infty} h_i(x + \lambda z, z) \quad (2.10)$$

Por otro lado, para  $z \neq 0$  en  $\mathbb{R}^n$ , tenemos dos alternativas para una restricción  $g_i$  arbitraria:  $z$  es dirección de  $g_i$  o no es.

i) Supongamos que  $g'_{i_{\infty}} \leq 0$ . Entonces, para  $\bar{x}$  un punto viable del problema (P) y  $\lambda > 0$  tenemos, por la Observación 2.25,

$$h'_i(\bar{x} + \lambda z, z) = P'_i(g_i(\bar{x} + \lambda z) / \lambda_k, \mu_i^k)[g'_i(\bar{x} + \lambda z, z)] / \lambda_k.$$

Cuando  $\lambda \rightarrow \infty$  tenemos dos alternativas:

- i)  $g_i(\bar{x} + \lambda z) \rightarrow -\infty$ . Entonces  $P'_i(g_i(\bar{x} + \lambda z) / \lambda_k, \mu_i^k) \rightarrow 0$  y  $g'_i(\bar{x} + \lambda z)$  es limitada, o
- ii)  $g_i(\bar{x} + \lambda z)$  es limitada. Luego  $P'_i(g_i(\bar{x} + \lambda z) / \lambda_k, \mu_i^k)$  es limitada y como  $g_i$  es convexa  $g'_i(\bar{x} + \lambda z) \rightarrow 0$ .

En ambos casos,

$$h'_i(\bar{x} + \lambda z, z) \rightarrow 0 \quad \text{cuando} \quad \lambda \rightarrow \infty. \quad (2.11)$$

ii) Se  $g'_{i_\infty} > 0$ ,  $g_i(\bar{x} + \lambda z) \rightarrow \infty$  y  $P'_i(g_i(\bar{x} + \lambda z)/\lambda_k, \mu_i^k) \rightarrow \infty$  cuando  $\lambda \rightarrow \infty$ .

O sea,

$$h'_i(\bar{x} + \lambda z, z) \rightarrow \infty. \quad (2.12)$$

Sustituyendo estos resultados en (2.10) y usando la Observación 2.23, tenemos:

$$L'_\infty(z, \mu^k, \lambda_k) = \begin{cases} f'_\infty(z) & \text{si } z \text{ es dirección de recesión de las funciones } g_i \\ \infty & \text{caso contrario.} \end{cases} \quad (2.13)$$

Finalmente, supongamos que existe  $z^0 \neq 0$  dirección de recesión de  $L_{\mu^k, \lambda_k}$ , luego, de (2.13),  $f'_\infty(z^0) \leq 0$  (o sea,  $z^0$  es dirección de recesión de las funciones  $g_i$  y de  $f$ ), lo que contradice la Observación 2.21.

Por tanto,  $L'_\infty(z, \mu^k, \lambda_k) > 0 \quad \forall z \neq 0 \text{ en } \mathbb{R}^n$ .

## BIBLIOGRAFÍA

- [1] Bertsekas, D.P., *Constrained Optimization and Lagrange Multiplier Methods*, Academic Press, New York, 1982.
- [2] Gonzaga, C. y Castillo, R., "Métodos de Lagrangeano Aumentado usando Penalidades Generalizadas para Programação não linear" Tesis, COPPE, UFRJ, 1998.
- [3] Hiriart-Urruty, J.-B. y Lemarechal, C., *Convex Analysis and Minimization Algorithms I*, 1 ed. New York, Springer-Verlag, 1993.
- [4] Hiriart-Urruty, J.-B. y Lemarechal, C., *Convex Analysis and Minimization Algorithms II*, 1 ed. New York, Springer-Verlag, 1993.
- [5] Hestenes, M., "Multiplier and Gradient Methods", *JOTA*, vol. 4, pp.303-320, 1969.
- [6] Iusem, A., *Augmented Lagrangian methods and proximal point methods for convex optimization*, Minicurso I workshop en Optimización, Florianópolis, Brasil, Diciembre 1997.

- [7] *Polyak, R. Y Teboulle, M., "Nonlinear rescaling and proximal-like methods in convex programming", 76, pp. 697-739, 1995.*
- [8] *Powell, M., "A method for nonlinear constraints in minimizations problems", in Optimización, R. Fletcher, Ed., Academic Pres, N.Y, pp.283-298, 1969.*
- [9] *Rockafellar, R. T., Convex Analisis, 1 ed. Princeton University Press, Princeton, N.J.1970.*
- [10] *Rockafellar R. T., "Augmented Lagrangians and applications of the proximal point algorithm in convex programming", Mathematics of Operations Researchs 1, pp. 97-116, 1976.*

# Películas de óxidos mixtos de cobre y tungsteno obtenidas por sol-gel: caracterización estructural y evaluadas como sensor de vapor

---

*Angélica Damián Briones, Yovan Rodríguez Daga, José Solís, Walter Estrada\**

## RESUMEN

En este trabajo se fabricó películas de óxido de tungsteno, con adición de Cu (II), mediante un proceso combinado de las técnicas de sol gel y rociado pirolítico. Para la obtención del gel isopolitúngstico se partió de una solución acuosa de tungstato de sodio dihidratado, la cual se pasó por una resina de intercambio iónico fuertemente ácida hasta obtener un pH aproximado de 1,1; para obtener el heteropolitunstato de cobre (II) se añadió al gel isopolitúngstico una solución acuosa de sulfato de cobre penta hidratado a diferentes concentraciones a fin de obtener una relación molar Cu/W entre 0,01 y 1,00. Con cada una de estas concentraciones se procedió la fabricación de las películas.

El parámetro más crítico para la obtención de las películas fue la temperatura, cuyo valor óptimo fue 220 °C. Posteriormente las películas fueron sometidos a un tratamiento térmico a 600 °C por 3 horas.

---

\* Laboratorio de Películas Delgadas, Escuela Profesional de Física, Facultad de Ciencias, Universidad Nacional de Ingeniería.

La estructura por difracción de rayos-x de la película delgada del óxido de tungsteno es la fase monoclinica del  $\text{WO}_3$ . A medida que se fue introduciendo Cu (II) al  $\text{WO}_3$  se observó la formación de una nueva fase, el  $\text{CuWO}_{4-x}$ , la misma que se hizo evidente a partir de la relación molar 0,10 de Cu/W. Para una relación molar Cu/W de 1,00 se observó únicamente la presencia de la fase  $\text{CuWO}_{4-x}$ .

Estas películas se usaron como sensor de vapor de etanol y butanol, notándose que la adición de Cu (II) al  $\text{WO}_3$  mejora el sensor debido probablemente a un incremento de la rugosidad y eventualmente de la porosidad de las películas con el incremento de Cu. La adición de Cu (II) permitió mejorar aproximadamente en 8 veces la sensibilidad al butanol respecto al  $\text{WO}_3$ , para una relación molar Cu/W de 0,10. Para el caso del etanol, la mayor sensibilidad se obtuvo para una relación molar Cu/W de 0,07.

## Introducción

La necesidad de dispositivos destinados a controlar atmósferas que pueden contener gases tóxicos o inflamables es cada vez más importante debido a una mayor conciencia de la protección del medio ambiente y la seguridad del ser humano<sup>1</sup>. Un elemento imprescindible para este propósito es el sensor de gas, dispositivo que detecta y cuantifica la cantidad de determinado gas en el ambiente que lo rodea. Los sensores de gas semiconductores se basan en óxidos metálicos semiconductores tales como son del  $\text{SnO}_2$  y  $\text{ZnO}^2$ . La resistencia de estos óxidos semiconductores cambia reversiblemente debido a la presencia de determinado gas o vapor en el ambiente.

Los sensores de gases semiconductores son usados como alarmas de fuga de gas natural, de incendios, sistemas de ventilación automática, etc. La selectividad y sensibilidad de los mismos se puede mejorar usando dopantes y aditivos<sup>3,4</sup>. Shaver<sup>5</sup> demostró que la conductividad de películas delgadas de  $\text{WO}_3$  activado con Pt cambia cuando se cambia la concentración de  $\text{H}_2$  en el medio que lo rodea. Los sensores de óxidos semiconductores operan a temperaturas entre 300 y 400 °C. J.L Solis et al.<sup>6</sup> reportaron que sensores de gas basados en nanopartículas de  $\text{WO}_3$  pueden operar a temperatura ambiente con una gran sensibilidad al  $\text{H}_2\text{S}$ , lo cual hace al  $\text{WO}_3$  un material muy interesante para esta aplicación.

La química del proceso sol-gel esta basado en la hidrólisis y condensación del precursor molecular, donde primeramente un sistema condensado, llamado sol, se transforma mediante un proceso de polimeramente continua a un gel. Esta polimerización esta basado en la formación de los polianiones cuya síntesis nos da

materiales con propiedades específicas en diversas áreas como la medicina, sensores, etc.<sup>7,8,9</sup>. Películas delgadas de  $\text{WO}_3$  sensibles al  $\text{O}_3$  han sido obtenidas por sol gel con precursores orgánicos<sup>10</sup>. Esta técnica es atractiva porque de una forma versátil, barata y simple se pueden obtener materiales nano-porosos, incrementando la difusión e interacción de las moléculas gaseosas con el material semiconductor. El gel obtenido en el proceso de sol-gel se puede utilizar como la solución a pulverizar por la técnica de rociado pirolítico, esta variante permite la obtención de películas delgadas muy porosas<sup>11</sup>.

El presente trabajo esta principalmente centrado en la optimización del óxido de tungsteno obtenido por la técnica de sol gel – rociado pirolítico para la detección de vapores de etanol y butanol, lo cual se ha logrado obtener mediante el ingreso sistemático de iones  $\text{Cu(II)}$  obteniéndose una mezcla del  $\text{WO}_3$  y  $\text{CuWO}_{4-x}$ .

## Procedimiento experimental

Las películas delgadas de  $\text{WO}_3$  con y sin adición de  $\text{Cu(II)}$  han sido preparadas por la técnica de sol gel – rociado pirolítico sobre substratos de vidrio y alúmina. El proceso de sol gel convencional se muestra en la figura 1(a)<sup>6</sup>. La idea fundamental de este proceso es de reemplazar las técnicas clásicas de alta temperatura por un proceso que se realice a temperaturas menores y en el cual sea posible formar una red vidriosa mediante polimerización de compuestos apropiados (monómeros) a baja temperatura.

El proceso de rociado pirolítico, consiste simplemente en pulverizar una solución sobre un substrato caliente<sup>12</sup>.

El esquema presentado en la figura 1(b) muestra el sistema de rociado pirolítico empleado, que consiste básicamente en: una cámara de vidrio, que sirve de colector de la solución y también contiene al pulverizador; por donde sale la solución. Encima de la cámara está una tobera de vidrio por donde ascienden las gotas más pequeñas hasta llegar a un substrato caliente, sobre el cual se produce la reacción pirolítica.

Al combinar ambos métodos nos da una nueva técnica que la denominamos sol gel – rociado pirolítico. En este proceso de rociado pirolítico utiliza como solución el sol obtenido por el proceso sol gel, produciéndose la gelificación y evaporación (proceso 6 en la fig. 1) del solvente sobre el substrato caliente.

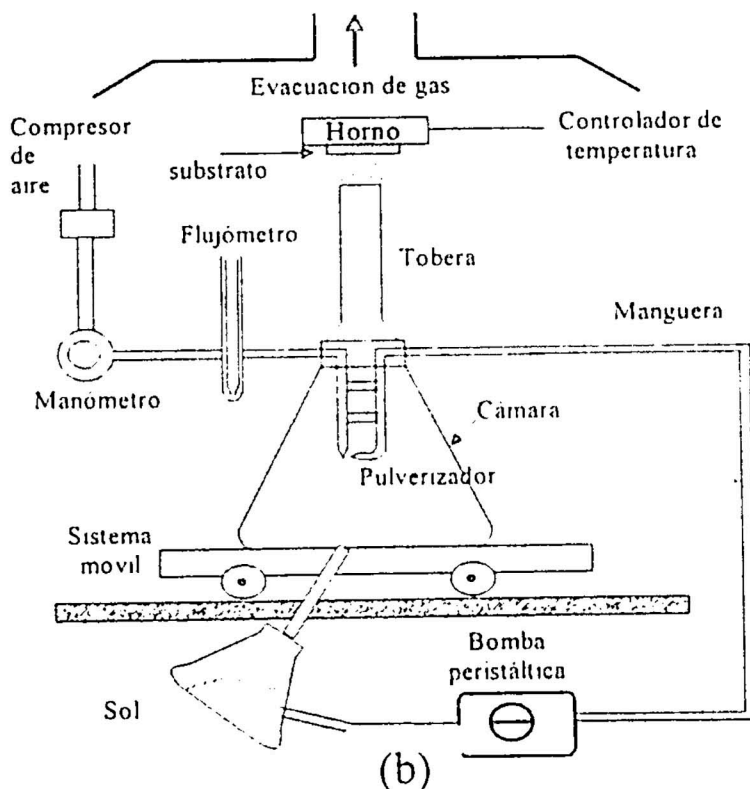
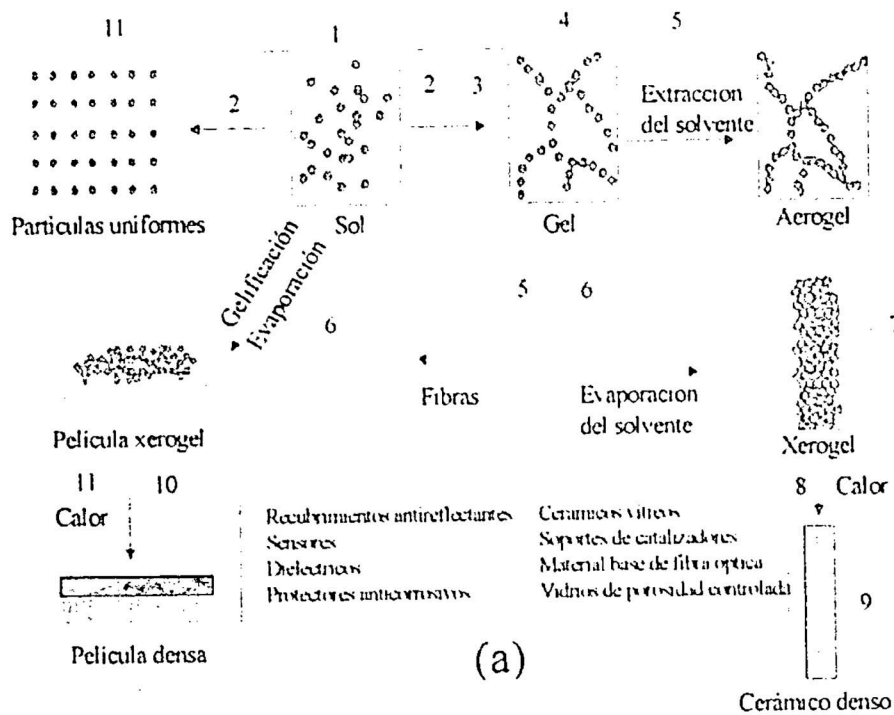


Figura 1. Esquema del proceso sol gel<sup>6</sup> (a), equipo de rociado pirolítico<sup>14</sup> (b).

Para la preparación de las películas delgadas de  $\text{WO}_3$  y la mezcla del  $\text{WO}_3$  con  $\text{CuWO}_{4-x}$  se sigue el diagrama de flujo mostrado en la figura 2<sup>13</sup>.

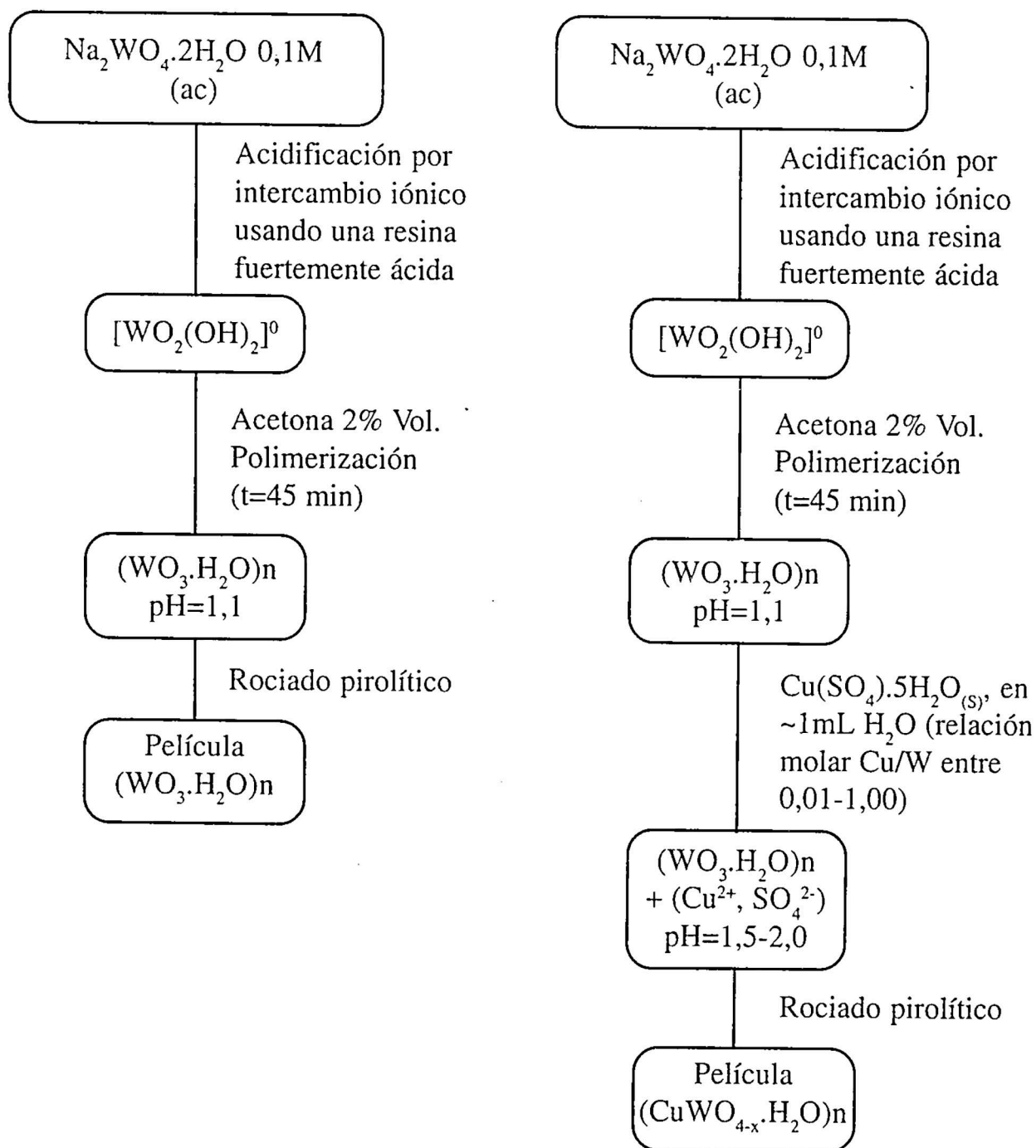


Figura 2. Diagrama esquemático de la obtención de películas de óxido tungsteno  $\text{WO}_3$  y óxidos mixtos de Cu y W por sol gel – rociado pirolítico.

Para iniciar el rociado pirolítico se debe esperar que la temperatura del sustrato se estabilice en  $220\text{ }^{\circ}\text{C}$ . El rociado se hizo durante 45 minutos.

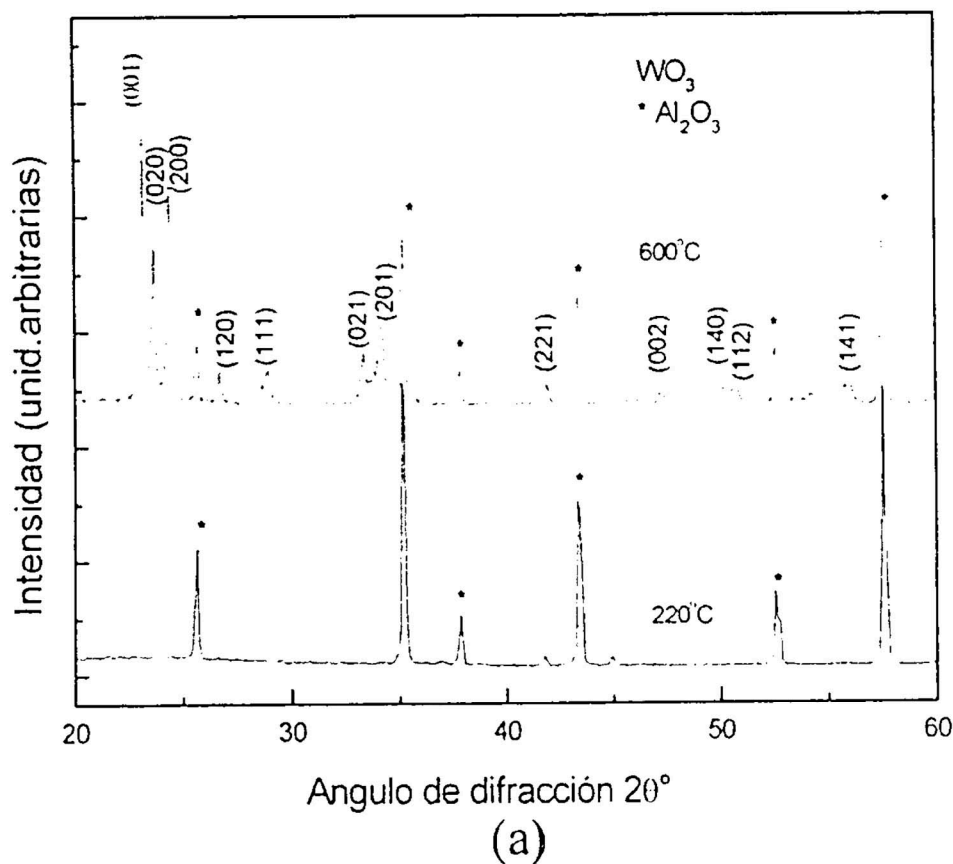


El estudio de las propiedades como sensor de gas de las películas delgadas se hizo utilizando un circuito eléctrico diseñado en la Facultad de Ciencias<sup>14</sup>.

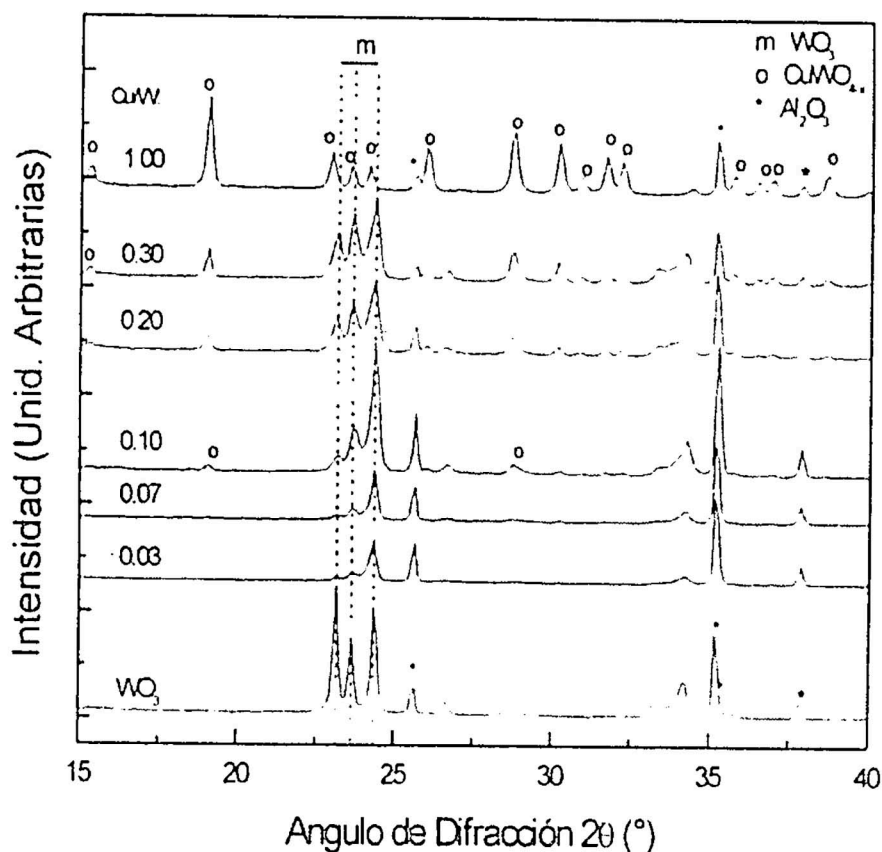
## Resultados y discusiones

### Caracterización estructural

El difractograma de la película delgada del óxido de tungsteno se muestra en la fig. 3(a) y en la fig 3 (b) se muestran el de las películas delgadas con adición de diferentes relaciones molares de Cu/W. La película delgada obtenida a 220 °C no presenta picos de difracción solo se observa los que corresponden a sustrato (fig. 3(a)), luego del tratamiento térmico sostenido a 600 °C, el óxido de tungsteno obtenido corresponde a la fase monoclinica del  $WO_3$  de acuerdo con los estándares<sup>15</sup> y los picos marcados con asterisco (\*) corresponde al sustrato ( $Al_2O_3$ ). Debido a que el sensor de gas semiconductor opera a temperaturas entre 300 y 500 °C y esta se cristaliza, entonces se procedió a realizar el mismo proceso a todas las películas obtenidas, se obtuvo a 220 °C y posteriormente se realiza un tratamiento térmico a 600 °C por 3 h.



En la figura 3(b) el difractograma para una relación molar Cu/W de 0 correspondiente al  $\text{WO}_3$  puro. Al añadir Cu (II) con una relación molar menor que 0,1 no presenta una fase nueva. Para relaciones molares mayores de 0,1 se observa la aparición de una nueva fase correspondiente al  $\text{CuWO}_{4-x}$ <sup>16</sup>. Al añadir Cu (II) con la relación molar de uno a uno con el tungsteno (Cu / W igual a 1), se obtiene únicamente la fase  $\text{CuWO}_{4-x}$ , como se muestra en la fig. 3(b).



(b)

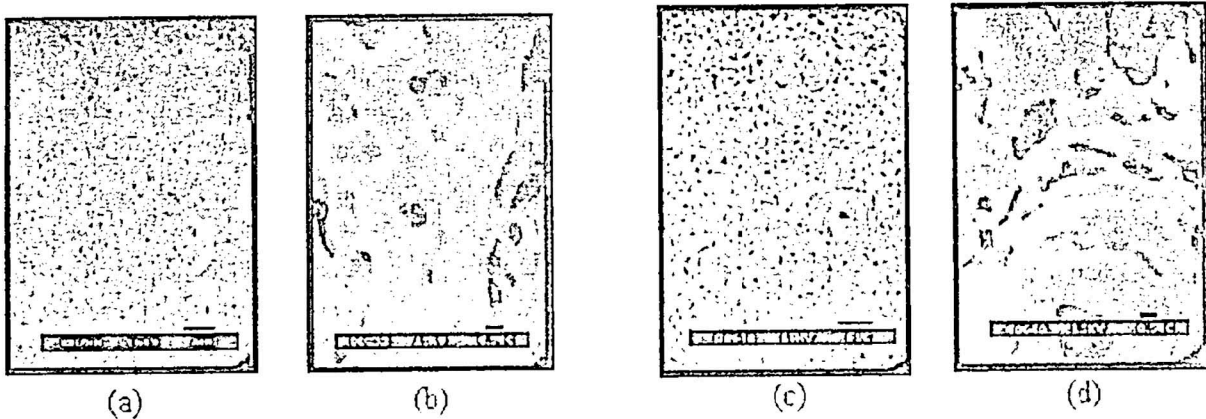
**Figura 3.** Difractograma de rayos-X (a) de la película delgada  $\text{WO}_3$  obtenida a 220 °C y con un posterior tratamiento a 600 °C por 3 h. y (b) el  $\text{WO}_3$  con adición de Cu (II) con relaciones molares Cu/W entre 0 y 1, obtenidas a 220 °C y con un posterior tratamiento a 600 °C.

### Caracterización morfológica

El estudio morfológico de las películas delgadas se realizó por microscopía electrónica de barrido, utilizando un microscopio Hitachi S500. En la figura 4(a) y (b) se observan las micrografías de las películas delgadas de  $\text{WO}_3$  obtenidas a 220 °C, las cuales están compuestas por hilos cuyas paredes tienen apariencia lisa. Posteriormente del tratamiento térmico a 600 °C (figura 4(c) y (d)), se observa que la morfología de la película cambia y que tienden a formar argollas semi-redondeadas

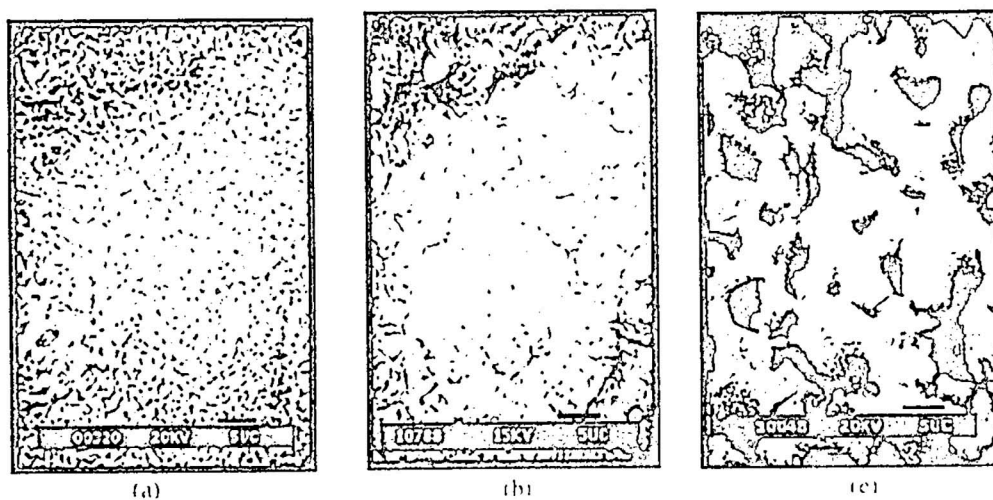
en cuyas paredes se empiezan a desarrollar aglomeraciones<sup>17</sup>, el cambio observado del material implica un reacomodo microestructural el cual esta relacionando a la cristalización del material a la fase monoclinica que se observo con la técnica de difracción de rayos X.

Baja magnificación: — 5 $\mu$ m, Alta magnificación: — 0,5 $\mu$ m



**Figura 4.** Micrografías SEM de las películas delgadas de óxido de tungsteno obtenidas a 220 °C (a) y (b), y después del tratamiento térmico a 600 °C (c) y (d).

La influencia en la morfología de la película delgada debido a la adición de Cu (II) al  $\text{WO}_3$  se muestra en la figura 5. Se observa que existe un incremento de la porosidad de la película hasta que la relación molar Cu/W es 0,10 (figura 5 (a), 5 (b)). Para relaciones molares mayores Cu/W 1,00 (figura 5 (c)), se observa que los poros tienden a aglomerarse, volviéndose mas compactas. La morfología de la fase  $\text{CuWO}_{4-x}$ , (figura 5(c)) muestra que no hay la porosidad observada en la fig, 5 (a) y 5 (b), la morfología que presenta es accidentada.



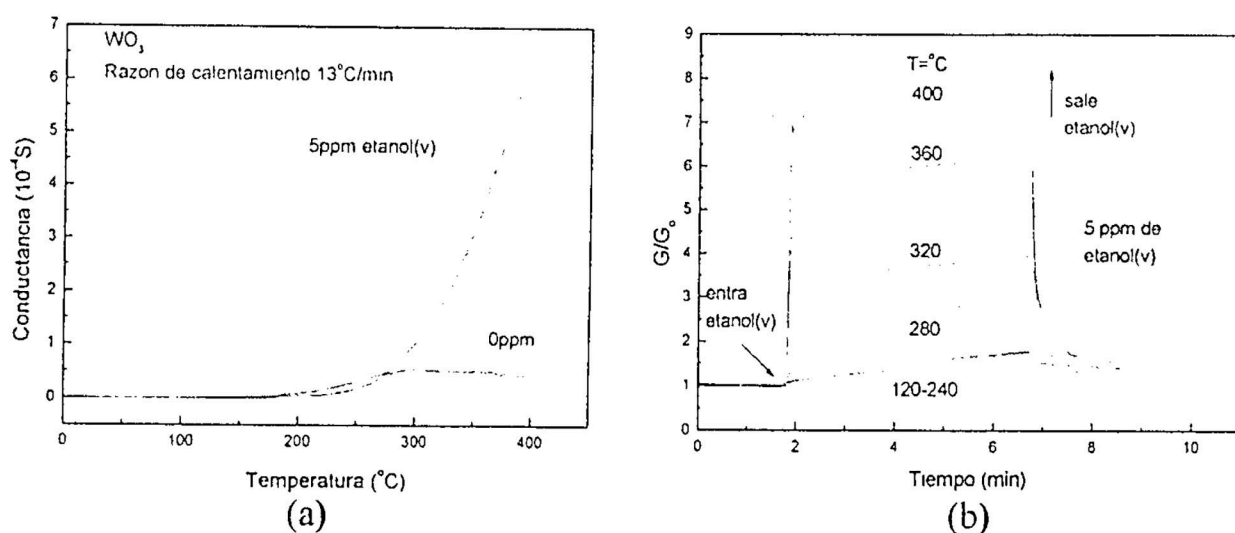
**Figura 5.** Micrografías SEM de baja magnificación (5 $\mu$ ) de las películas delgadas de  $\text{WO}_3$  con adición de cobre (II) obtenidas a 220 °C, y con tratamiento térmico a 600 °C en relación molar Cu/W: 0,03 (a), 0,10 (b), 1,00 (c).

## Evaluación de la películas como sensor de vapor de alcoholes

La caracterización eléctrica del sensor para la detección de vapor de etanol y butanol se realizó con las películas de  $\text{WO}_3$  y  $\text{WO}_3$  con adición de Cu (II) para diferentes relaciones molares de Cu/W. Todas las películas delgadas se obtuvieron bajo las mismas condiciones utilizando la técnica de sol gel -rociado pirolítico.

Para estudiar las propiedades del óxido de tungsteno con y sin adición de Cu (II) expuesto al vapor de etanol y butanol, se varió la concentración del vapor y la temperatura de operación.

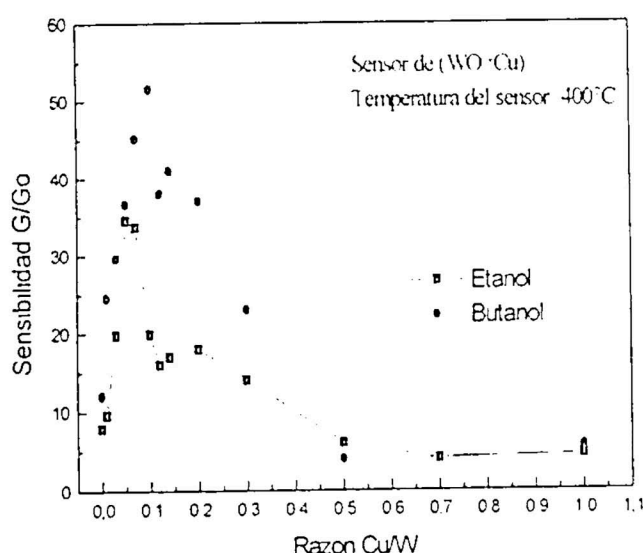
Un parámetro importante en la operación de los sensores semiconductores es la temperatura de operación. En la figura 6(a) se muestra el cambio de conductancia de la película delgada de  $\text{WO}_3$  para una temperatura de operación variable entre 25 °C y 400 °C con y sin exposición a 5 ppm de etanol. Se observa que a bajas temperaturas la conductancia con y sin exposición al etanol es la misma, pero a partir de 200 °C la conductancia de la película expuesta a etanol es mayor respecto a la película que no esta expuesta al etanol. Esta diferencia se incrementa conforme aumenta la temperatura de operación. En la figura 6 (b) se observa la variación de la conductancia normalizada,  $G(t)/G_0$ , donde  $G(t)$  es la conductancia del sensor en función del tiempo y  $G_0$  es la conductancia antes que se ingrese etanol a la cámara. Se tiene que la temperatura donde la película tiene una mejor respuesta ocurre cuando la temperatura de operación del sensor es 400 °C.



**Figura 6.** (a) Conductancia de la película delgada del  $\text{WO}_3$  a temperaturas entre 25 °C y 400 °C en una atmósfera sin y con 5 ppm de etanol y (b) respuesta de la conductancia normalizada  $G(t)/G_0$  del  $\text{WO}_3$  a 5 ppm de vapor de etanol en aire a temperaturas entre 120 °C y 400 °C.

La sensibilidad de un sensor de gas se define como el cociente entre la conductancia  $G$ , del sensor expuesto al gas medida después de 2 min. que el gas esta en la cámara de prueba y la conductancia del sensor expuesto solo al aire  $G_0$ ,  $S=G/G_0$ .

La adición de Cu (II) al  $WO_3$  aumenta la sensibilidad al etanol y butanol para bajas proporciones como se muestra en la figura 7. Para el caso del etanol se alcanza una sensibilidad de 35 para películas delgadas obtenidas con una relación molar Cu/W entre 0,05 y 0,07; para películas obtenidas con mayores relaciones molares Cu/W decrece la sensibilidad al etanol. Para el butanol se observa algo similar que para el etanol, para este caso una sensibilidad de 50 con películas delgadas obtenidas con una relación molar Cu/W de 0,10; luego la sensibilidad decrece para películas obtenidas con mayores relaciones molares Cu/W. La disminución de la sensibilidad estaría relacionada por la formación de una nueva fase,  $CuWO_{4-x}$ , determinado por difracción de rayos X el cual no tendría buenas propiedades como sensor y la disminución de la porosidad de la morfología de la película.



**Figura 7.** Sensibilidad de las películas delgadas con diferentes razones molares Cu/W, entre 0 y 1 para una concentración de 5 ppm de vapor de etanol y butanol, a una temperatura de operación de  $400^\circ C$ .

## Conclusiones

Se pudieron encontrar parámetros óptimos para la obtención de películas delgadas de  $WO_3$  y  $CuWO_{4-x}$  sobre alúmina. La temperatura de obtención de las películas delgadas es  $220^\circ C$  y un tratamiento a  $600^\circ C$  por 3 horas. Con las técnicas de caracterización de difracción de rayos X nos permiten afirmar que la película depositada es óxido de tungsteno en fase monoclinica.

Con la adición de cobre (II) al sol isopolitúngsteno a partir de una razón molar Cu/W de 0,10 permite la formación del  $\text{CuWO}_{4-x}$ . La formación de éste compuesto alcanza cuando se tiene una relación atómica de un cobre por una de tungsteno, es decir Cu/W de 1.

Con respecto a sus propiedades como sensor, la adición de cobre al óxido de tungsteno permite mejorar la sensibilidad al vapor de etanol y butanol con respecto al óxido de tungsteno puro donde se alcanza mejorar la sensibilidad en 8 veces en comparación al óxido de tungsteno puro y luego decae. Las películas estudiadas muestran mayor sensibilidad al butanol en comparación al etanol. La adición del cobre (II) da como resultado un aumento de respuesta como sensor de vapores de alcoholes, la cual esta relacionada con la porosidad de las películas, permitiendo que el vapor tenga una mayor área. Mientras que para relaciones molares altas la porosidad disminuye, lo cual también esta relaciona con la formación de la nueva fase  $\text{CuWO}_{4-x}$ .

## Agradecimientos

Este trabajo ha sido realizado gracias al apoyo de la Facultad de Ciencias de la UNI, del International Science Program de la Universidad de Uppsala en Suecia y al CONCYTEC.

## BIBLIOGRAFÍA

- [1] *J.L. Solis, Estudio de las propiedades como sensor de gas del  $\text{SnWO}_4$ , Tesis de Maestría, Facultad de Ciencias, Universidad Nacional de Ingeniería, Lima, Perú, 1995.*
- [2] *T. Seiyama, A. Kato, K. Fujishi, and M. Nagatanui, Anal. Chem. 34 (1962) 1502.*
- [3] *H. Nanto, H. Sokooshi, T. Kawai, and T. Usuda, J. Mater. Sci. Lett. 11 (1992) 235.*
- [4] *H. Nanto, H. Sokooshi, T. Kawai, and T. Usuda, Sensors and Actuators 13-14 (1992) 715.*
- [5] *P.J. Shaver, Appl. Phys. Lett. 11 (1967) 255.*
- [6] *J.L. Solis, S. Saukko, L.B. Kish, C.G. Granqvist, and V. Lantto, Sensors and Actuators B 77 (2001) 316.*
- [7] *J.C. Brinker and G.W. Sherer, Sol - Gel Science, Academic Press, San Diego (1990).*
- [8] *M.T. Pope, Heteropoly and Isopoly Oxometalates, Springer-Verlag Berlin, Heidelberg, 1983.*
- [9] *F. Cotton and G. Wilkinson, Química Inorgánica Avanzada, Limusa S.A., México (1993).*

- [10] C. Cantalini, M.Z. Atashbar, Y. Li, M.K. Ghantasala, S. Santucci, W. Wlodarski, and M. Passacatando, *J. Vac. Sci. Technol. A* 17(4) (1999) 1973.
- [11] A. Medina, *Síntesis y Caracterización de Películas de Óxido de Tungsteno con Adición de Fósforo y Estaño. Aplicación como Materiales Electrocrómicos*, Tesis de Licenciatura, Facultad de Ciencias, Universidad Nacional de Ingeniería, Lima, Perú, 1998.
- [12] F. Paraguay, *Obtención de películas delgadas de Óxido de Zinc, Caracterización Estructural y Óptica*, Tesis de Licenciatura, Facultad de Ciencias, Universidad Nacional de Ingeniería, Lima, Perú, 1993.
- [13] A. Medina, *Síntesis y Caracterización de películas de Óxido de Tungsteno con adición de Fósforo y Estaño. Aplicaciones como Materiales Electro crómicos*. Tesis de Licenciatura Facultad de Ciencias, Universidad Nacional de Ingeniería, Lima, Perú, 1998.
- [14] J. Morales, *Estudio de las películas de ZnO:X como detector de vapor de etanol*, Tesis de Licenciatura, Facultad de Ciencias, Universidad Nacional de Ingeniería, Lima, Perú, 2000.
- [15] *International Centre for Diffraction Data. Powder Diffraction Files*, 1997.
- [16] E. Geber and L. Kihlborg, *Acta Chem. Scand.*, 21 (1967) 2575.
- [17] J. Aguilar, *Tesis de Licenciatura, Facultad de Ciencias, Universidad Nacional de Ingeniería, Lima, Perú, (será publicado)*.

# Funciones de Green para un problema del electromagnetismo

---

*Carlos Enrique Valcárcel Flores\**

## *RESUMEN*

Estudiamos las ecuaciones de Maxwell dentro de un espacio  $\Omega$  y las resolvemos usando el método de la función de Green. Encontramos que el resultado es el mismo que se obtiene por el denominado método de las imágenes. La ventaja está en que aquí resolvemos las ecuaciones sin hipótesis ad-hoc, solo considerando las condiciones iniciales y de contorno del problema.

## *ABSTRACT*

We study the Maxwell's equations within an space we will called  $\Omega$  and we resolve them using the method of the Green's function. We find that our result is the same we get by the method of the images. The principal advantage is that here we resolve the equations without ad-hoc hypothesis, we only consider the full well-posed problem, i.e. initial conditions plus boundary conditions of the problem.

---

\* Grupo estudiantil de Física Teórica. K19981125@uni.edu.pe



## Introducción

Consideramos una región del espacio  $\Omega$  dentro de la cual existe un medio homogéneo con constantes  $\epsilon_0$  y  $\mu_0$ , además de una distribución de cargas  $\rho(\vec{r}, t)$  y una distribución de corrientes  $\vec{J}(\vec{r}, t)$ . Ahora convertiremos las ecuaciones de Maxwell a las ecuaciones de D'Alembert para el potencial eléctrico y el potencial vectorial magnético. Luego resolveremos las ecuaciones de D'Alembert por medio de las funciones de Green, considerando todas las soluciones, esto es, el potencial **retardado** y el potencial **avanzado** junto con las condiciones de contorno.

Dentro de  $\Omega$  el campo eléctrico  $\vec{E}$  y el campo magnético  $\vec{B}$  vienen determinados por las ecuaciones de Maxwell

$$\nabla \cdot \vec{E} = \frac{\rho}{\epsilon_0} \quad (1)$$

$$\nabla \cdot \vec{B} = 0 \quad (2)$$

$$\nabla \times \vec{E} + \partial \vec{B} = \vec{0} \quad (3)$$

$$\nabla \times \vec{B} - \frac{1}{c^2} \partial \vec{E} = \mu_0 \vec{J}, \quad (4)$$

donde estamos usando la notación  $\partial \equiv \partial / \partial t$ . La ecuación (2) nos permite definir el **potencial vectorial magnético**  $\vec{A}$  según

$$\vec{B} \equiv \nabla \times \vec{A}. \quad (5)$$

Reemplazamos (5) en (3) y podemos definir el **potencial escalar eléctrico**  $\varphi$  según

$$\vec{E} \equiv -\nabla \varphi - \partial \vec{A} \quad (6)$$

Al hecho de definir los potenciales también se le denomina "resolver" la segunda y tercera ecuación de Maxwell. Reemplazamos (6) en (1) y obtenemos

$$\frac{1}{c^2} \partial^2 \varphi - \nabla^2 \varphi = \frac{\rho}{\epsilon_0} + \partial(\nabla \cdot \vec{A} + \frac{1}{c} \partial \varphi) \quad (7)$$

Además, usando la identidad vectorial

$$\nabla \times \nabla \times \vec{A} = \nabla(\nabla \cdot \vec{A}) - \nabla^2 \vec{A} \quad (8)$$

en la ecuación (4) nos lleva a una nueva ecuación para el potencial vectorial magnético

$$\frac{1}{c^2} \partial^2 \vec{A} - \nabla^2 \vec{A} = \mu_0 \vec{J} - \nabla(\nabla \cdot \vec{A} + \frac{1}{c^2} \partial \varphi) \quad (9)$$

Sabemos que existe una indeterminación en las definiciones dadas en (5) y (6), para esto definamos ahora un nuevo campo  $\eta$ , tal que, a partir de él formamos nuevos potenciales de la forma

$$\vec{A}' = \vec{A} + \nabla \eta \quad (10)$$

$$\varphi' = \varphi - \partial \eta \quad (11)$$

Un simple cálculo nos lleva a verificar que

$$\vec{B} = \nabla \times \vec{A}' \quad (12)$$

$$\vec{E} = -\nabla \varphi' - \partial \vec{A}' \quad (13)$$

Esto quiere decir que  $\vec{E}$  y  $\vec{B}$  son invariantes ante una **transformación de escala**. Por lo tanto existe la libertad de poder escoger los potenciales de manera conveniente. Escogemos entonces una restricción conocida como el **gauge de Lorentz** o **calibre de Lorentz**

$$\nabla \cdot \vec{A} + \frac{1}{c} \partial \varphi = 0 \quad (14)$$

que reemplazándola en las ecuaciones (7) y (9) nos permite obtener ecuaciones simples para los campos. Debemos tener presente que el gauge de Lorentz no es el único "calibre" útil en electromagnetismo, existe también el gauge de Coulomb  $\nabla \cdot \vec{A} = 0$ .

Veamos como afectan estos nuevos potenciales (10) y (11) en las ecuaciones de onda (7) y (9)

$$\frac{1}{c^2} \partial^2 \varphi' - \nabla^2 \varphi' = \frac{1}{c^2} \partial^2 \varphi - \nabla^2 \varphi - \partial \left( \frac{1}{c^2} \partial^2 \eta - \nabla^2 \eta \right) \quad (15)$$

$$\frac{1}{c^2} \partial^2 \vec{A}' - \nabla^2 \vec{A}' = \frac{1}{c^2} \partial^2 \vec{A} - \nabla^2 \vec{A} + \nabla \left( \frac{1}{c^2} \partial^2 \eta - \nabla^2 \eta \right) \quad (16)$$

Además escogemos  $\eta$  tal que

$$\frac{1}{c^2} \partial^2 \eta - \nabla^2 \eta = 0 \quad (17)$$

De esta manera resultan invariantes las ecuaciones de onda (o lo que es lo mismo, se mantiene el **principio de la covariancia**).

## La ecuación de D'Alembert

Las ecuaciones para los potenciales, bajo el gauge de Lorentz son

$$\frac{1}{c^2} \partial^2 \varphi - \nabla^2 \varphi = \frac{\rho}{\epsilon_0} \quad (18)$$

$$\frac{1}{c^2} \partial^2 \vec{A} - \nabla^2 \vec{A} = \mu_0 \vec{J} \quad (19)$$

Tanto la ecuación para el potencial eléctrico como las ecuaciones de las tres componentes del potencial vectorial magnético tienen la forma

$$\frac{1}{c^2} \partial^2 \Psi - \nabla^2 \Psi = f \quad (20)$$

que es denominada **ecuación de D'Alembert**.

Sabemos que la **función de Green**  $G(x^\mu) = G(t, x, y, z)$ , es solución de

$$\frac{1}{c^2} \partial^2 G - \nabla^2 G = \delta(\vec{r} - \vec{r}_0) \delta(t - t_0) \quad (21)$$

Esto es, la función de Green es la respuesta a la acción de una fuente puntual localizada en  $\vec{r} = \vec{r}_0$  y que actúa en el instante  $t = t_0$ ; la función de Green nos brinda la información sobre el efecto de este impulso mientras este se aleja del punto  $\vec{r}_0$ .

### Proposición [1]

Si  $G = G(x^\mu)$  satisface (21) entonces

$$G(\vec{r}, \vec{r}_0; t, t_0) = G(\vec{R}, r) \quad (22)$$

donde  $\vec{R} \equiv \vec{r} - \vec{r}_0$  y  $\tau \equiv t - t_0$ .

Ver referencias [2] y [3].

En tal caso las funciones de Green que satisfacen (21) son

$$G^{\pm}(\vec{R}, \tau) = -\frac{1}{4\pi} \frac{\delta(\tau \pm \frac{|\vec{R}|}{c})}{|\vec{R}|} \quad (23)$$

a la función  $G^+$  se le denomina **función de Green adelantada** y a  $G^-$  se le denomina **función de Green retardada**.

Tenemos las ecuaciones siguientes

$$\frac{1}{c^2} \partial^2 \Psi - \nabla^2 \Psi = f \quad (24)$$

$$\frac{1}{c^2} \partial^2 G - \nabla^2 G = \delta(\vec{r} - \vec{r}_0) \delta(t - t_0) \quad (25)$$

Multiplicamos la primera ecuación por  $G$ , la segunda por  $\Psi$  y luego restamos e integramos

$$\Psi = \int_{\Omega} G f d^3 \vec{r}_0 + T \quad (26)$$

donde el primer término es la solución inhomogénea y

$$T \equiv -\frac{1}{c^2} \int_{\Omega} (f \partial_0 G - G \partial_0 f) d^3 \vec{r}_0 + \int_{\partial \Omega} (G \nabla_0 \Psi - \Psi \nabla_0 G) d\vec{S}_0 dt_0 \quad (27)$$

es el término debido a las condiciones de contorno.

## Forma de los potenciales

Comparando las ecuaciones (20) y (26) con (18) y (19), tenemos la forma del potencial eléctrico

$$\varphi = \frac{1}{4\pi\epsilon_0} \int_{\Omega} \frac{\rho(\vec{r}_0, t - \frac{|\vec{r} - \vec{r}_0|}{c})}{|\vec{r} - \vec{r}_0|} d^3 \vec{r}_0 + \frac{1}{4\pi\epsilon_0} \int_{\Omega} \frac{\rho(\vec{r}_0, t + \frac{|\vec{r} - \vec{r}_0|}{c})}{|\vec{r} - \vec{r}_0|} d^3 \vec{r}_0 + T \quad (28)$$

Obtenemos también la expresión para el potencial vectorial magnético

$$\bar{A} = \frac{\mu_0}{4\pi} \int_{\Omega} \frac{\bar{J}(\bar{r}_0, t - \frac{|\bar{r} - \bar{r}_0|}{c})}{|\bar{r} - \bar{r}_0|} d^3\bar{r}_0 + \frac{\mu_0}{4\pi} \int_{\Omega} \frac{\bar{J}(\bar{r}_0, t + \frac{|\bar{r} - \bar{r}_0|}{c})}{|\bar{r} - \bar{r}_0|} d^3\bar{r}_0 + T \quad (29)$$

El primer término en la solución de  $\varphi$  y  $\bar{A}$  es el denominado **potencial retardado** y el segundo término es el **potencial avanzado**.

## Aplicaciones

Veremos que bajo condiciones iniciales y de contorno apropiadas, podemos transformar a las ecuaciones (28) y (29) a la forma que usualmente se encuentran en los textos de electromagnetismo.

Consideremos como condiciones iniciales  $G(\bar{r}, t; \bar{r}_0, t_0)$  y  $\partial G$  nulas para  $t < t_0$ . Esto quiere decir que no existe impulso antes de  $t_0$ . Esto anula el término del potencial avanzado. Además tomaremos como espacio  $\Omega = R^3$  y

$$\lim_{\bar{r} \rightarrow \infty} \psi(\bar{r}, t) = 0 \quad (30)$$

con  $\psi = \varphi, \bar{A}$ . Luego la solución para los potenciales toma la forma

$$\varphi = \frac{1}{4\pi\epsilon_0} \int_{\Omega} \frac{\rho(\bar{r}_0, t - \frac{|\bar{r} - \bar{r}_0|}{c})}{|\bar{r} - \bar{r}_0|} d^3\bar{r}_0 \quad (31)$$

$$\bar{A} = \frac{\mu_0}{4\pi} \int_{\Omega} \frac{\bar{J}(\bar{r}_0, t - \frac{|\bar{r} - \bar{r}_0|}{c})}{|\bar{r} - \bar{r}_0|} d^3\bar{r}_0 \quad (32)$$

Para poder entender mejor el uso de los potenciales adelantados, consideremos el caso unidimensional (onda que viajan en el eje X) para el cual la función de Green (integrando en  $x_0$  e  $y_0$ ) es de la forma

$$G^{\pm}(x, x_0, t, t_0) = \frac{c}{2} \left[ 1 \pm \Theta \left( \frac{|x - x_0|}{c} - \tau \right) \right] \quad (33)$$

Existe una carga puntual localizada en  $x = a$  y  $t = 0$  ( $\rho(x,t) = q\delta(x - a) \delta(t)$ ) y una pared en  $x = 0$ , esto nos lleva a proponer condiciones iniciales y de contorno de la forma

$$\partial\varphi(t=0) = 0 \quad (34)$$

$$\lim_{x \rightarrow \infty} \varphi = \varphi(x=0) = 0 \quad (35)$$

Obtenemos el término de contorno  $T$  como

$$T = \frac{1}{2c} \varphi(x - ct, 0) + \frac{1}{2c} \varphi(x + ct, 0) \quad (36)$$

y el término de la solución inhomogénea es de la forma

$$\frac{cq}{2} \left[ 1 - \Theta \left( \frac{|x - x_0|}{c} - t \right) \right] \quad (37)$$

Como se muestra en las dos últimas ecuaciones tenemos tanto potenciales retardados como avanzados, el potencial retardado está presente al momento de iniciarse las ondas (debido a que la solución inhomogénea representa el efecto de las fuentes) mientras que el potencial avanzado aparece debido al término de contorno en la pared.

## Conclusión

La solución de la ecuación de D'Alembert para los potenciales permiten tanto la existencia de potenciales retardados como avanzados. El hecho de eliminar alguna solución (por ejemplo el potencial avanzado) va de acuerdo con las condiciones de contorno.

## BIBLIOGRAFÍA

- [1] *W. Hauser, Introduction to the principles of Electromagnetism. Addison-Wesley, 1971.*
- [2] *E. Butkov, Mathematical Physics. Addison-Wesley, Mass., 1968.*
- [3] *P. M. Morse y H. Feshbach, Methods of the Theoretical Physics. McGraw-Hill.*

# Cristalografía Cósmica

---

Armando Bernui Leo \*

## RESUMEN

Un problema importante en Cosmología es la búsqueda de las propiedades globales del universo, y una de las más importantes es la múltiple conexidad. Si nuestro universo tuviese esa propiedad, ella se manifestaría con la presencia en el cielo de imágenes múltiples de un dado objeto cósmico, tal como una galaxia. La Cristalografía Cósmica es un método estadístico que permite descubrir dichas propiedades conociendo las coordenadas espaciales ( $r$ ,  $\theta$ ,  $\varphi$ ) de todos los integrantes de una dada clase de objetos cósmicos. Aquí mostramos las virtudes y defectos de este método y como aprovecharlo –según los datos observacionales disponibles– para descubrir las propiedades globales del universo.

## ABSTRACT

An important open problem in Cosmology is the search for topological properties of the universe, e.g. multiple connectedness which manifest itself through the presence in the sky of multiple images of a given cosmic object like a galaxy. Given a distribution of objects in a 3-space, the method of Cosmic Crystallography (CC) aims to discover their topological properties looking for distance correlations in Pair Separation Histograms (PSH). Since these correlations appear only when the 3-space containing the objects is multiply connected, their presence in the histogram evinces the existence of multiple images caused by topological properties. Here we present an introductory review of this method.

---

\* Facultad de Ciencias, Universidad Nacional de Ingeniería. [bernui@uni.edu.pe](mailto:bernui@uni.edu.pe)

## Introducción

Un problema que continua a desafiar al ser humano es determinar si vivimos en un universo de tamaño finito o si por el contrario el universo es infinito<sup>1</sup>. Desde la antigüedad muchos pensadores se ocuparon de esto, algunos creyeron llegar a conclusiones definitivas en uno u otro sentido basándose en criterios que hoy reconocemos como carentes de base científica. Actualmente estamos **casi** como al principio pues no tenemos argumentos teóricos ni observacionales que inclinen la balanza hacia uno u otro lado.

Las propiedades globales de un espacio, también llamadas propiedades topológicas, proporcionan información sobre su forma, y entre estas propiedades tal vez la más importante para la Cosmología es la *múltiple conexidad*<sup>2</sup>. Esta propiedad da origen a una clase interesante de 3-espacios, aquellos múltiplemente conexos, sin frontera y con volumen finito, denominados *espacios compactos*. Así por ejemplo, sabemos que existen 6 espacios tri-dimensionales con geometría Euclideana, orientables y compactos.

Los espacios múltiplemente conexos son conocidos desde hace mucho tiempo [1, 2, 3, 4], sin embargo la posibilidad de considerarlos como candidatos del universo se inicia recién con el modelo denominado *pequeño universo* de Ellis & Schreiber [5]. En este modelo el universo es representado por un 3-espacio *compacto*, cuyo volumen finito se repite varias veces hasta alcanzar el horizonte observable. Por datos recientes sobre –las anisotropías de– la radiación cósmica de fondo sabemos que el tamaño de este tri-espacio compacto no puede ser muy pequeño comparado con el volumen encerrado por el horizonte observable [6, 7].

Sin embargo, la posibilidad que nuestro universo sea pequeño no está totalmente descartada ni teórica ni observacionalmente (en este último caso solo se ha limitado su tamaño mínimo). Desde el punto de vista teórico los modelos cosmológicos de Friedmann [8] y Lemaître [9], describen bien la evolución de la materia y la radiación después de la Última Superficie de Dispersión (USD) y son soluciones de las ecuaciones de Einstein-Hilbert (EH). No obstante, es bien sabido que cualquier solución de la ecuación diferencial EH es válida localmente, i.e. en una vecindad del espaciotiempo. Consecuentemente, hipotizar acerca de las

---

<sup>1</sup> denominamos universo al espacio tri-dimensional (en adelante 3-espacio) donde se ubican los objetos cósmicos, tales como estrellas, galaxias, aglomerados de galaxias, radiación cósmica de fondo, etc.

<sup>2</sup> un espacio es simplemente conexo cuando toda curva cerrada es contraíble a un punto; un espacio es múltiplemente conexo cuando existe al menos una curva cerrada que no es contraíble a un punto.



propiedades de homogeneidad e isotropía del espaciotiempo sólo puede referirse al sentido *local* de tales propiedades, y no al sentido *global*.

De hecho, la geometría local en los modelos cosmológicos de Friedmann-Lemaître (FL) satisface las propiedades de homogeneidad e isotropía local descrita por la métrica de Robertson-Walker (RW) [10]  $ds^2 = dt^2 - a^2(t)d\sigma^2$ , donde  $a(t)$  es el factor de escala, y  $d\sigma^2 = d\chi^2 + F_k^2(\chi) [d\theta^2 + \text{Sen}^2 \theta d\varphi^2]$  es la métrica de los 3-espacios  $t = \text{constante}$  y que denotaremos por  $\mathcal{M}_k$ . Las funciones  $F_k(\chi) \equiv \text{Sen } \chi$ ,  $\chi$ ,  $\text{Senh} \chi$ , donde  $k = +1, 0, -1$  es el parámetro de curvatura normalizado caracteriza la geometría esférica, Euclideana e hiperbólica respectivamente. Más aún, sabemos que los  $\mathcal{M}_k$  son espacios de curvatura constante obtenidos de la partición (o teselaje) del 3-espacio  $\tilde{\mathcal{M}}_k$  de la siguiente forma  $\mathcal{M}_k = \tilde{\mathcal{M}}_k / \Gamma$  donde  $\Gamma$  es un grupo discreto de isometrías sin punto fijo<sup>3</sup> del espacio  $\tilde{\mathcal{M}}_k$ , el cual es llamado el *recubrimiento universal* de  $\mathcal{M}_k$  (e.g. [11]).

Aquí remarcamos que la elección del parámetro  $k = +1, 0, -1$  en la métrica de RW determina sólo la geometría local de  $\mathcal{M}_k$  (y de paso también la de  $\tilde{\mathcal{M}}_k$ ). Desde que el espacio simplemente conexo  $\tilde{\mathcal{M}}_k$  y el espacio múltiplemente conexo  $\mathcal{M}_k$  comparten las propiedades *locales* de homogeneidad e isotropía podemos dotar indistintamente ambos espacios con la geometría local de RW. Consecuentemente la relatividad general (o cualquier otra teoría métrica de la gravitación) no puede distinguir entre  $\tilde{\mathcal{M}}_k$  o  $\mathcal{M}_k$ , tampoco puede determinar cuál es la estructura *global* – también denominada *estructura a grande escala*– del universo.

En el campo *observacional* notamos que aún es controversial la escala correcta a la cual la materia se distribuye homogéneamente. Hoy en día, las observaciones más profundas indican un universo con regiones extremadamente densas –los llamados super-clusters de galaxias, como por ejemplo “la gran muralla” y también estructuras tipo filamento– pero también grandes vacíos [12].

Frente a este panorama surgen entonces algunas preguntas interesantes: ¿cuáles son las propiedades globales del universo? ¿cómo podemos usar los datos observacionales para determinarlas? ¿el universo es simple o múltiplemente conexo?

En este artículo introductorio sólo consideramos el caso de los 3-espacios con geometría Euclideana  $\tilde{\mathcal{M}}_0 = \mathcal{R}^3$ . En la sección 2 haremos una breve introducción al método de la Cristalografía Cósmica y a los histogramas de separación por pares

<sup>3</sup> esto garantiza que el espacio obtenido por el teselaje no tiene puntos privilegiados (no existe un “centro”).

(HSP). En la sección 3 deduciremos la expresión analítica o esperada teóricamente del HSP, denominada Histograma de Separación por Pares Esperado (HSPE). En la sección 4 discutimos las probabilidades analíticas o teóricas de las isometrías Euclidianas y calculamos explícitamente el valor correspondiente a las translaciones de Clifford. Luego en la sección 5 mostramos, en histogramas simulados numéricamente, los resultados de las diferentes isometrías Euclidianas en los HSP. Finalmente en la última sección, hacemos un resumen de lo estudiado y sugerimos una alternativa al problema de la inexistencia de catálogos observacionales completos; este tipo de catálogos resultan necesarios para el método de la Cristalografía Cósmica, al menos en su formulación original.

## El método de la Cristalografía Cósmica

La Cristalografía Cósmica (CC) es un método estadístico que busca descubrir las propiedades globales del universo [13-24]. El método se basa en el hecho que la distribución de objetos en un 3-espacio se relaciona con sus propiedades geométricas y topológicas. El método busca *correlaciones de distancia* entre dichos objetos a través de *Histogramas de Separación por Pares* (HSP). Para esto es necesario que los objetos cuyas separaciones se graficarán en el histograma sean de un mismo tipo, que estén distribuidos en todo el universo observable y por supuesto que se conozcan sus posiciones  $(r, \theta, \varphi)$  en el 3-espacio. En tal caso los objetos pasan a formar parte de un *catálogo completo*.

Las hipótesis en que se basa el método de la CC son:

- Un espacio 3-dimensional con una geometría local, es decir una métrica, la cual sirve para calcular distancias entre los objetos cósmicos localizados en dicho espacio y listados en el catálogo. Aquí asumimos que la métrica del 3-espacio  $\tilde{\mathcal{M}}_k = \mathcal{R}^3$ , es

$$d\sigma^2 = dr^2 + r^2 [d\theta^2 + \sin^2\theta d\varphi^2] = dx^2 + dy^2 + dz^2.$$

- Un catálogo completo  $C$  de  $N$  fuentes cósmicas –con coordenadas  $(r_i, \theta_i, \varphi_i)$ ,  $i = 1, 2, \dots, N$ – las cuales se distribuyen siguiendo alguna ley de distribución. Aquí usaremos catálogos de objetos cósmicos puntuales, “aleatoriamente” generados y homogéneamente distribuidos. Es decir

$$\rho(r, \theta, \varphi) = \text{cte.}$$

- Un horizonte (o *cut-off*), es decir

$$r_i \leq R_H$$

para  $i = 1, 2, \dots, N$ <sup>4</sup>.

De hecho, el resultado que se evidencia en los HSP's es notable, como veremos en la sección 5. Si los puntos en estudio están distribuidos en un espacio múltiplemente conexo  $\mathcal{M}_k$ , entonces los histogramas muestran picos o pequeñas desviaciones respecto de una curva base correspondiente a una distribución de puntos en el espacio simplemente conexo  $\tilde{\mathcal{M}}_k$ . Estos picos o las pequeñas desviaciones ponen de manifiesto las correlaciones entre pares de puntos, las cuales son esperadas debido a la existencia –en la distribución de puntos– de imágenes múltiples producidas por las diferentes isométricas de  $\mathcal{M}_k$ <sup>5</sup>. Los correlaciones tipo picos corresponden a translaciones puras –también llamadas translaciones de Clifford–, y para las otras isometrías el HSP muestra deformaciones apreciables respecto de la curva base que corresponde al caso simplemente conexo [19-22].

De hecho cuando los puntos se ubican en un espacio simplemente conexo  $\mathcal{M}_k = \tilde{\mathcal{M}}_k = S^3, \mathcal{R}^3, \mathcal{H}^3$ , estas *correlaciones* están ausentes. Más aún la posición, ancho, y amplitud relativa de las correlaciones son un efecto característico de cada una de las isometrías. Por esta razón es de fundamental importancia conocer el HSPE para poder identificar las pequeñas desviaciones y diferenciarlas del ruido estadístico siempre presente en los HSP. Conceptos como HSP, HSPE, Histograma de Separación por Pares Medio (HSPM), etc. han sido explícitamente desarrollados en [22].

Como consecuencia de la múltiple conexidad, cualquier 3-espacio con esta propiedad cuyo dominio fundamental<sup>6</sup> se localize al menos parcialmente dentro del horizonte presenta necesariamente imágenes múltiples. Así, si el universo es múltiplemente conexo entonces existen imágenes de objetos cósmicos que son, en principio, observables.

En resumen, si contásemos con catálogos completos de objetos cósmicos observados, estos contendrían información sobre las isometrías del 3-espacio, las cuales se pondrían en evidencia en los HSPM [13-23].

<sup>4</sup>  $R_H \sim 10,000 h^{-1} \text{ Mpc}$

<sup>5</sup> Las isometrías forman el grupo discreto  $\Gamma$  (sin punto fijo) del 3-espacio  $\tilde{\mathcal{M}}_k$ , cada una de ellas genera diferentes imágenes múltiples de cada punto excepto cuando  $\Gamma = \text{identidad}$ , pues en tal caso  $\mathcal{M}_k = \tilde{\mathcal{M}}_k$  es decir se trata del 3-espacio simplemente conexo.

<sup>6</sup> los espacios múltiplemente conexos se representan por un poliedro donde cada par de caras está identificado; este poliedro recibe el nombre de dominio fundamental.

## Histograma de Separación por Pares

Ahora estudiaremos como obtener, desde su definición, los histogramas antes mencionados. Un HSP es un histograma normalizado del número de pares de objetos vs. la distancia entre ellos (más generalmente: vs. cualquier función de la distancia). Pasaremos a construir formalmente los HSP.

Consideremos un 3-espacio  $\mathcal{M}_k$  dotado con una geometría local de signatura  $k$ . Sea  $\mathcal{B}_R \subset \tilde{\mathcal{M}}_k$  una vecindad de radio  $R$  centrada en el origen de coordenadas  $O$  y conteniendo objetos cósmicos. Denotamos por  $\mathcal{B}_R$  el universo observable, siendo  $R \equiv R_{\text{USD}}$  el radio del universo observable. Una regla de selección bien definida es necesaria para homogenizar los objetos cósmicos de un catálogo, esta puede ser por ejemplo la *luminosidad* del objeto o la presencia de una cierta línea espectral característica en su espectro observado. Un catálogo *completo*  $C$  es una lista homogénea de todos los objetos cósmicos presentes en el universo observable los cuales satisfacen la regla de selección dada y tienen coordenadas  $(r, \theta, \varphi)$ , donde  $r$  es la distancia radial al origen  $r \in (0, 2R]$

Dado el catálogo  $C$  denotamos por  $\phi(\ell)$  el número de pares de objetos separados por una distancia  $\ell \in (0, 2R]$ . Ahora dividamos el intervalo  $(0, 2R]$  en  $m$  sub-intervalos o *bins* de igual longitud  $\Delta\ell = 2R/m$ , donde cada sub-intervalo tiene entonces la forma  $J_i = (\ell_i - (\Delta\ell/2), \ell_i + (\Delta\ell/2)]$ ,  $i = 1, 2, \dots, m$ , con centro en  $\ell_i = (i - 1/2) \Delta\ell$ . Por lo tanto

$$\Phi(\ell_i) = \frac{2}{N(N-1)\Delta\ell} \sum_{\ell \in J_i} \phi(\ell), \quad (1)$$

es realmente un conteo normalizado de la cantidad de pares de objetos separados por una distancia  $\ell_i$  que yace en el sub-intervalo  $J_i$ . Podemos observar que  $\Phi(\ell_i)$  es en efecto un conteo normalizado  $\sum_{i=1}^m \Phi(\ell_i) \Delta\ell = 1$ ; de esta manera podremos comparar histogramas hechos con catálogos conteniendo diferente cantidad de objetos. Más aún, si bien el histograma es realmente el gráfico de la función  $\Phi$  vs.  $\ell_i$  es usual referirnos a la función  $\Phi(\ell_i)$  como siendo el HSP.

Vamos ahora a obtener el HSPE para un 3-espacio simplemente conexo  $\tilde{\mathcal{M}}_k$ . El HSPE viene definido por

$$\Phi_{\text{esperado}}(\ell_i) \equiv \frac{1}{\Delta\ell} \int_{J_i} \mathcal{P}(\ell) d\ell, \quad (2)$$

donde  $\mathcal{P}(\ell)$  es la densidad de probabilidad que un par arbitrario de objetos  $p$ ,

$q \in \mathcal{B}_R \subset \tilde{\mathcal{M}}_k$  del catálogo  $\mathcal{C}$  —donde  $x_p, x_q$  son las coordenadas de dichos objetos— esté separado por una distancia  $\ell$ , es

$$\mathcal{P}(\ell) = \int_{\mathcal{B}_R} \int_{\mathcal{B}_R} d^3x_p \rho(x_p) d^3x_q \rho(x_q) \delta(d_k(x_p, x_q) - \ell), \quad (3)$$

donde  $\delta$  es la delta de Dirac,  $d_k(x_p, x_q)$  es la función distancia entre  $p$  y  $q$ , la cual claramente depende de la geometría local (i.e.  $k = -1, 0, +1$ ). Notemos que  $\rho(x_p) d^3x_p$  es la probabilidad de encontrar el objeto  $p$  en el volumen  $d^3x_p$ . La probabilidad total de encontrar  $p$  en todo  $\mathcal{B}_R$  debe ser naturalmente uno, esto es  $\int_{\mathcal{B}_R} \rho(x_p) d^3x_p = 1$ . Así mismo podemos también verificar que la densidad de probabilidad  $\mathcal{P}(\ell)$  está normalizada

$$\int_0^{2R} \mathcal{P}(\ell) d\ell = 1. \quad (4)$$

Para una distribución homogénea de objetos en  $\mathcal{B}_R \in \mathcal{R}^3$  encontramos

$$\rho(p) = \Theta(R - d(O, x_p)) \frac{3}{4\pi R^3}, \quad (5)$$

donde  $O$  es el origen de coordenadas y  $\Theta$  es la función de Heaviside  $\Theta(x) = (1 + x|x^{-1})/2$ ,  $x \in \mathcal{R}$ . La expresión  $\mathcal{P}(\ell)$  ha sido obtenida en [24] para el caso de los 3-espacios simplemente conexos  $\tilde{\mathcal{M}}_0 = \mathcal{R}^3, \mathcal{S}^3, \mathcal{H}^3$ , y para una distribución homogénea de objetos —ec. (5)— correspondiente a catálogos completos. Para  $\mathcal{R}^3$  obtuvimos

$$\mathcal{P}(\ell) = \frac{3\ell^2 \Theta(2R - \ell)}{16R^6} (4R + \ell) (2R - \ell)^2 \quad (6)$$

Si el intervalo  $J_i$  es suficientemente pequeño, una aproximación apropiada para el HSPE es

$$\Phi_{esperado}(\ell_i) \simeq \mathcal{P}(\ell_i). \quad (7)$$

Finalmente, por definición del HSPE observamos que el ruido en los HSP's está dado por

$$\eta(\ell_i) = \Phi(\ell_i) - \Phi_{esperado}(\ell_i). \quad (8)$$

Las figuras 1a y 1b ilustran muy bien este punto. La curva suave en la figura corresponde a la ec. (6).

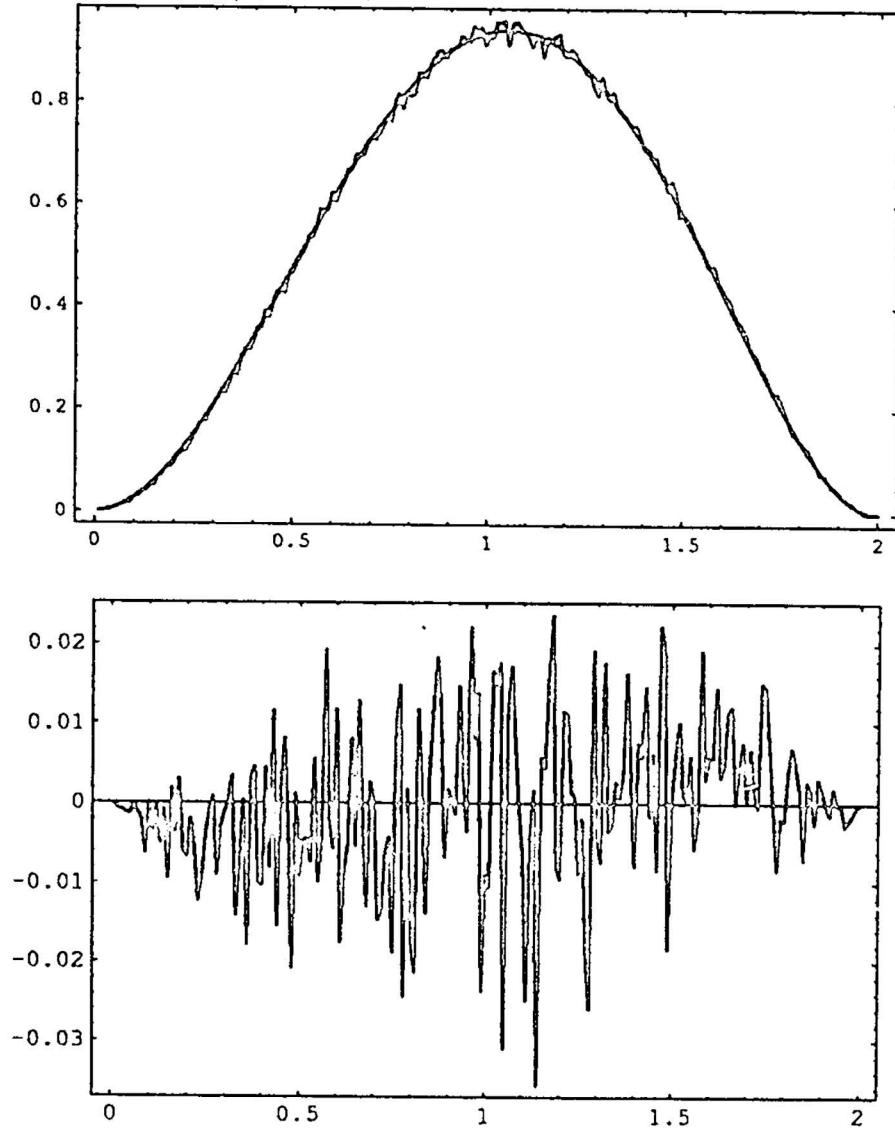


Figura 1. (a) HSPE de  $\mathcal{R}^3$  y HSPM para  $K = 50$  catálogos con  $N \sim 150$  objetos cada uno en el universo observable ( $R=1$ ). (b) Diferencia entre el HSPM y el HSPE graficados en (a).

### Isometrías en $\mathcal{R}^3$

Las isometrías son transformaciones de un espacio en él mismo que dejan invariante su métrica.

**Proposición 1.** Una traslación en  $\mathcal{R}^3$  es una función

$$\begin{aligned} T_L : \mathcal{R}^3 &\rightarrow \mathcal{R}^3, \\ x &\rightarrow x + L. \end{aligned} \quad (9)$$

Es posible mostrar que, el conjunto de traslaciones forma un grupo isomorfo a  $\mathcal{R}^3$  y que ellas son isometrías en  $\mathcal{R}^3$ .

**Definición.-** Una traslación que no depende del punto  $x$  se denomina traslación de Clifford.

**Corolario.-** Toda traslación en  $\mathcal{R}^n$  es una traslación de Clifford.

**Proposición 2.** Una transformación ortogonal es una función lineal  $\gamma : \mathcal{R}^3 \rightarrow \mathcal{R}^3$  que preserva el producto interno

$$\langle \gamma(x_1), \gamma(x_2) \rangle = \langle x_1, x_2 \rangle, \quad (10)$$

$\forall x_1, x_2 \in \mathcal{R}^3$ .

**Proposición 3.** Una matriz real  $3 \times 3$  se llama ortogonal si la transformación lineal asociada es ortogonal. Además una matriz  $R$  es ortogonal si  $R^T R = I$ .

**Proposición 4.** Una transformación ortogonal es una isometría Euclideana, sin embargo no es una traslación de Clifford.

**Definición.-** Se denota por  $O(n)$  al grupo de transformaciones ortogonales  $n$ -dimensionales, o también al grupo de matrices ortogonales  $n$ -dimensionales.

**Proposición 5.** Una función  $\gamma : \mathcal{R}^3 \rightarrow \mathcal{R}^3$  es una isometría en  $\mathcal{R}^3$  si y solo si  $\gamma = T_L$  o  $R$ , es decir  $\gamma(x) = Rx + L$ ,  $\forall x \in \mathcal{R}^3$  donde  $L \in \mathcal{R}^3$  y  $R \in O(3)$ . Denotaremos  $\gamma \equiv (R, T_L)$ .

Esta proposición afirma que toda isometría Euclideana puede ser descompuesta de manera única en una transformación ortogonal seguida de una traslación de Clifford.

**Proposición 6.** El conjunto de isometrías Euclidianas  $(R_i, T_{L_i})$  forma el grupo  $Isom(\mathcal{R}^n)$ .

Estudiemos ahora el caso más simple, es decir un 3-espacio que solo posee isometrías de traslación pura, esto es traslaciones de Clifford, nos referimos al tri-toro  $T^3$ . Dado un catálogo con objetos en el poliedro fundamental de  $T^3$  encontramos



que hay contribuciones a la densidad de probabilidad  $\mathcal{P}(\ell)$  que vienen de las isometrías de traslación pura y que se manifiestan en los HSP como picos agudos (con intensidad y en posiciones bien definidas).

La probabilidad que un par de objetos  $p, q \in \mathcal{B}_R$  esté correlacionado por la isometría  $g \in \Gamma$  viene dado por

$$\mathcal{P}^g(\ell) = \frac{1}{n} \int_{\mathcal{B}_R} \int_{\mathcal{B}_R} d^3x_p d^3x_q \rho(x_p) \rho(x_q) \delta(d_k(x_p, x_q) - \ell) \delta(x_q - g(x_p)) \quad (11)$$

$$= \frac{1}{n} \int_{\mathcal{B}_R} d^3x_p \rho(x_p) \rho(g(x_p)) \delta(d_k(x_p, g(x_p)) - \ell), \quad (12)$$

donde  $d_k(x_p, g(x_p))$  es la distancia entre el par  $(p, q) = (p, g(p))$  correlacionado por la isometría  $g$ ; como se trata de una traslación de Clifford tal distancia es independiente de  $p$ .

Así encontramos que la probabilidad que un par de objetos  $p, q \in \mathcal{B}_R \subset \mathcal{R}^3$ , i.e.  $k = 0$ , correlacionado por la isometría de traslación pura  $g \in \Gamma$  de longitud  $\lambda_g$ , es

$$\mathcal{P}^g(\ell) = \frac{v_g}{N} \delta(\ell - \lambda_g), \quad (13)$$

debido a la isometría de traslación  $g(x_p) = x_p + L_g$ , donde  $L_g \in \mathcal{R}^3$  y  $\lambda_g \equiv |L_g|$  es la longitud de dicha traslación. Luego  $d_0(x_p, g(x_p)) = \lambda_g$ ; además  $n \equiv N/V$ ,  $V = (4\pi/3) R^3$ ,  $N$  es el número de objetos en  $\mathcal{B}_R$ , y el coeficiente  $v_g$  viene dado por

$$v_g = 1 - \frac{3\lambda_g}{4R} + \frac{\lambda_g^3}{16R^3}. \quad (14)$$

Usando la ec. (13) en la definición (2) obtenemos

$$\Phi_{esperado}^g(\ell_i) \equiv \frac{1}{\Delta\ell} \int_{J_i} \mathcal{P}^g(\ell) d\ell = \frac{v_g}{\Delta\ell N} \delta_{i,\lambda_g}, \quad (15)$$

donde

$$\delta_{i,\lambda_g} = \begin{cases} 0 & \text{si } \lambda_g \notin J_i \\ 1 & \text{si } \lambda_g \in J_i \end{cases}, \quad (16)$$

y donde observamos que cuando la longitud  $\lambda_g$ , correspondiente a la traslación  $g \in \Gamma$ , coincide con el intervalo de distancias  $J_i$  entonces dicha isometría hace una



gran contribución al HSP (que como veremos se manifestará como un pico agudo), caso contrario no habrá contribución. Si representamos al  $\mathbb{R}^3$ -poliedro fundamental del tri-toro  $T^3$  como un cubo de lado  $L$ , entonces las isometrías de su grupo  $\Gamma$  (i.e. las traslaciones puras que permiten identificar sus caras opuestas, ver el cubo superior izquierdo de la figura 2) son  $\lambda_g = L$  (en la dirección de cada eje  $\pm x, \pm y, \pm z$ ),  $L\sqrt{2}$  (en las direcciones  $\pm xy, \pm xz, \pm yz$ ), ... etc.

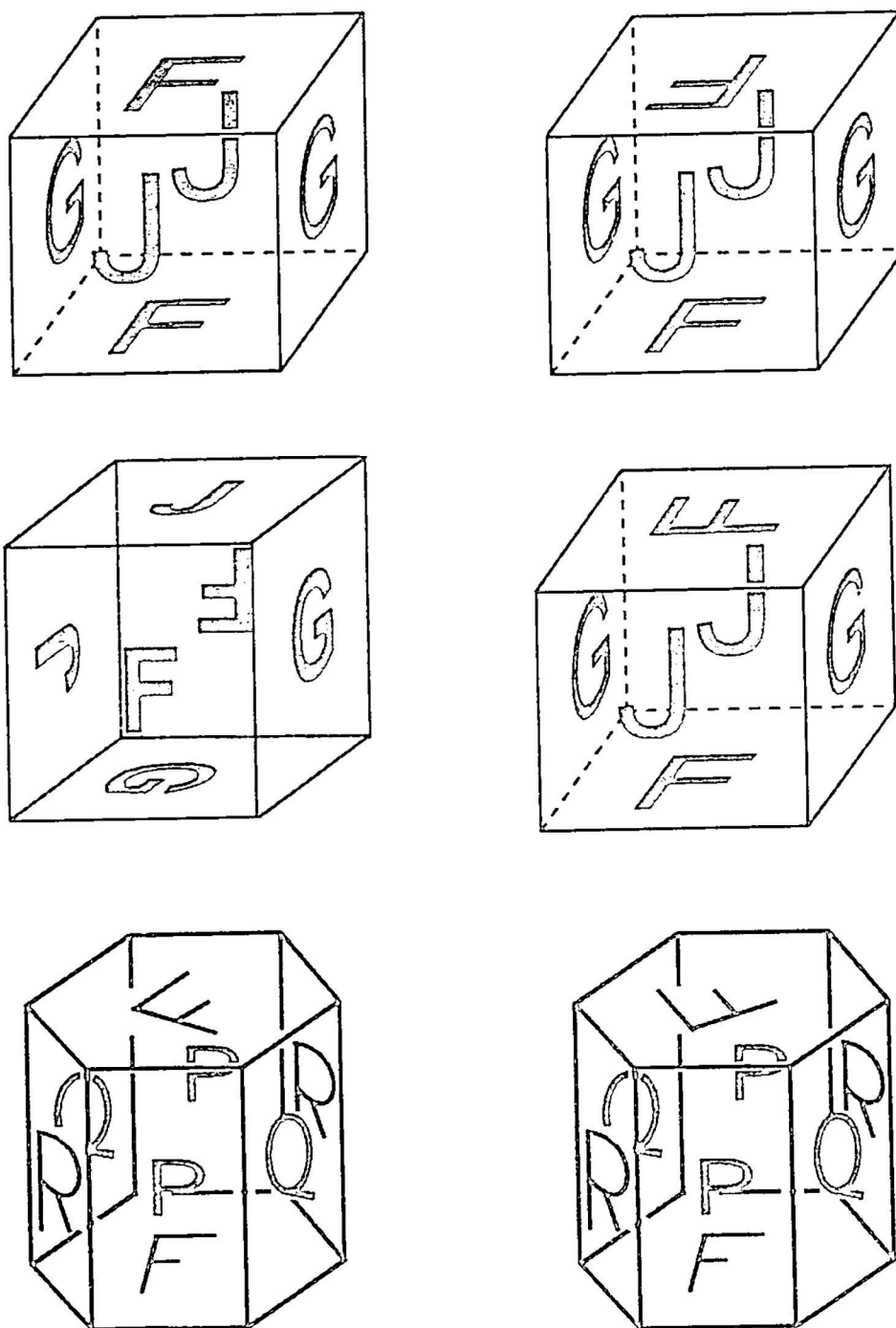


Figura 2. Los poliedros fundamentales de los seis 3-espacios compactos orientables Euclidianos; el cubo superior izquierdo es el tri-toro [3,4,25].

En los 3-espacios compactos orientables Euclidianos existen isometrías que no son traslaciones de Clifford, son las traslaciones seguidas de una rotación, denominadas *screw motion*<sup>7</sup>. Observemos los 3-espacios Euclidianos (compactos y orientables) cuyos poliedros fundamentales se muestran en la figura 2, escojamos el centro del poliedro como origen de coordenadas y un sistema de ejes cartesianos (e.g. perpendiculares a las caras de los poliedros). Notamos entonces que solo el tri-toro tiene sus 3 pares de caras opuestas identificados por sendas traslaciones de Clifford. En cambio, todos los otros 3-espacios (**excepto** el cubo ubicado debajo del tri-toro, en adelante denominado  $G_6$ ) presentan identificaciones entre caras que pueden descomponerse<sup>8</sup> como una rotación (o transformación ortogonal) alrededor de un eje que une los centros de las caras seguidas de una traslación pura a lo largo de ese mismo eje.

La isometría de screw motion  $g \in \Gamma$  se puede entonces representar como

$$g = \gamma_{i,\theta}^L = T_{L_i} \circ R_i(\theta), \quad (17)$$

donde  $R_i$  es una matriz que rota el 3-espacio un ángulo  $\theta$  alrededor del eje  $i$  (e.g.  $i = x, y, z$ ) y  $T_{L_i}$  es una traslación —de magnitud  $L_i$ — a lo largo del eje  $i$ . Consideremos por ejemplo un screw motion  $\gamma_{z,\alpha}^L$ , este se descompone en la rotación de un ángulo  $\alpha$  alrededor del eje  $z$

$$R_z(\theta) = \begin{pmatrix} \cos \alpha & \sin \alpha & 0 \\ -\sin \alpha & \cos \alpha & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad (18)$$

seguido de la traslación

$$T_{L_z}(x', y', z') = (x', y', z' + L). \quad (19)$$

$\forall (x', y', z') \in \mathcal{R}^3$ . De este modo, una traslación de Clifford es un caso particular de screw motion correspondiente a una rotación de 0 grados.

Veamos por ejemplo las isometrías para el poliedro que aparece al lado derecho del tri-toro (en adelante este 3-espacio lo denominamos  $\mathcal{T}_\pi$ )

$$g_1(x, y, z) = T_{L_x} = (x + L, y, z), \quad (20)$$

$$g_2(x, y, z) = T_{L_y} = (x, y + L, z), \quad (21)$$

$$g_3(x, y, z) = T_{L_z} \circ R_z(\pi/2) = (-x, -y, z + L), \quad (22)$$

<sup>7</sup> literalmente 'movimiento de sacacorchos': avanza y-rotación

<sup>8</sup> esto se debe al hecho que los 3-espacios son orientables

En cambio las isometrías para el 3-espacio  $G_6$  son

$$g_1(x, y, z) = (x + L, -y, -z) \quad (23)$$

$$g_2(x, y, z) = (-x, z + L, y) \quad (24)$$

$$g_3(x, y, z) = (-x, z, y + L) \quad (25)$$

las cuales no se pueden representar como las isometrías de *screw motion*, sino como rotaciones alrededor de alguna arista seguidas de alguna traslación (observar con cuidado  $G_6$  que es el cubo debajo del tri-toro en la figura 2).

## Simulaciones numéricas

Similarmente a los cálculos realizados con las traslaciones de Clifford podemos también calcular la probabilidad teórica de –que un par de objetos del catálogo esté correlacionado por– cualquier otra isometría Euclideana. Sin embargo, preferimos ser más didácticos e ilustrar gráficamente el efecto de las diferentes isometrías en los histogramas (HSPM).

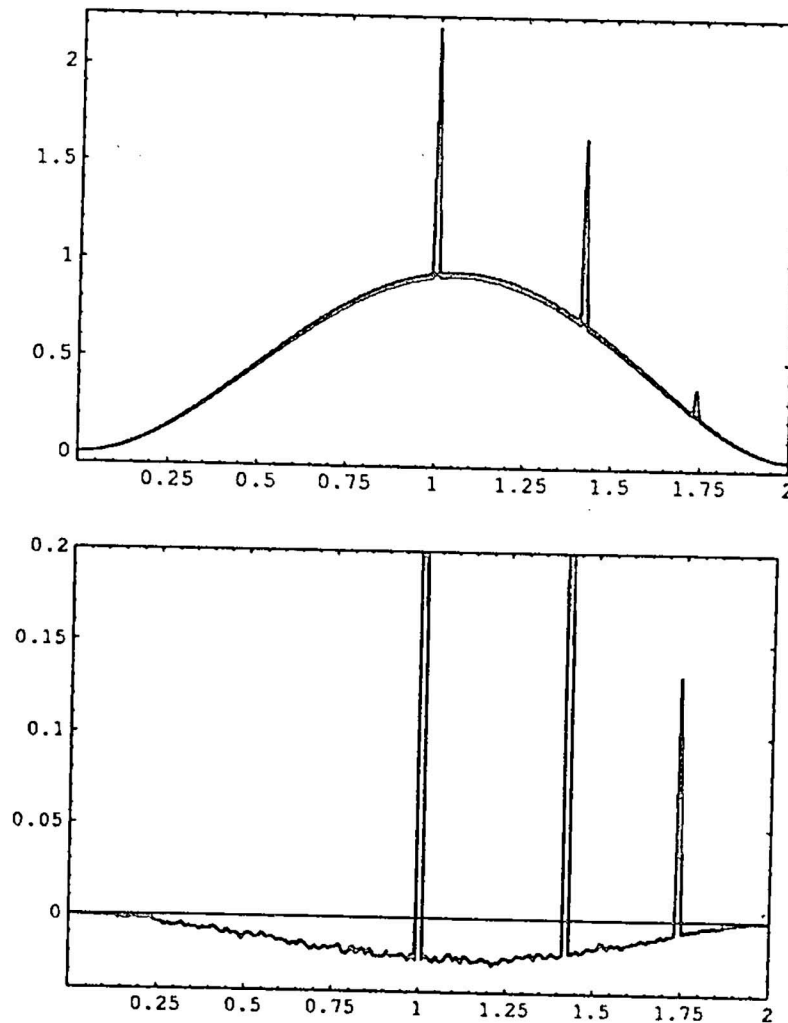
Estudiaremos ahora tres casos, es decir a los 3-espacios múltiplemente conexos  $T^3$ ,  $T_\pi$ ,  $G_6$ . Esto lo hacemos porque en ellos se presentan todas las isometrías Euclideanas de los 3-espacios orientables. Juntamente con el histograma de cada uno de estos 3-espacios dibujaremos el HSPM del 3-espacio simplemente conexo  $\mathcal{R}^3$  con el objetivo de mostrar que las isometrías aparecen en el primero pero no en el segundo de ellos.

Como vimos en la sección anterior, la contribución de las traslaciones de Clifford es tipo delta y localizada en valores discretos correspondientes a las isometrías traslacionales. Debido a esto ellas se manifiestan en los histogramas como picos agudos muy notorios, en cambio las otras isometrías (i.e. las no-traslacionales) se manifiestan en forma continua (o esparcida) y aparecen como pequeñas desviaciones respecto de la curva base. Además está el problema que en los histogramas siempre existe el ruido estadístico, así dichas isometrías no-traslacionales no son siempre evidentes. Debido a esto necesitamos una estrategia que haga visibles aún las señales topológicas más débiles presentes en el histograma. Nuestra estrategia es doble. Primero, dado que el ruido estadístico es aleatorio, consideraremos el promedio de histogramas, ya que al promediar  $K$  histogramas el ruido disminuye en un factor  $1/\sqrt{K}$ . Así, en todas las simulaciones que siguen hemos considerado el promedio de  $K = 10,000$  histogramas simulados. También en ellas escogimos  $m = 200$ ,  $\Delta\ell = 0,01$  y  $R = 1$  (excepto para el caso de  $G_6$  donde  $m = 141$  y  $R = \sqrt{2}/2$ ). Segundo, graficaremos junto con cada HSPM la diferencia entre el histograma medio

del espacio múltiplemente conexo y el del simplemente conexo ( $\mathcal{R}^3$ ). Como vemos, de esta manera sí conseguimos poner en evidencia las señales topológicas presentes, pero no siempre visibles, en los histogramas.

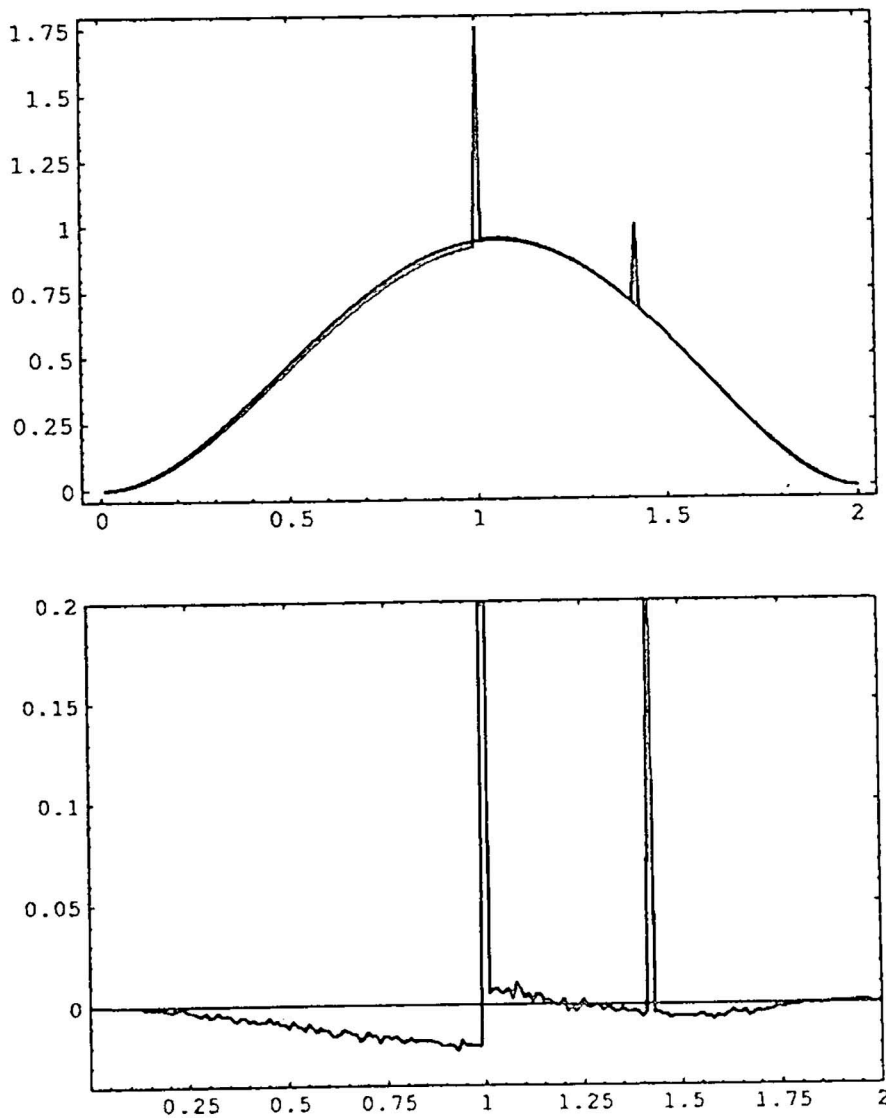
Vamos a mostrar ahora el efecto de las isometrías Euclideanas, comenzaremos con el caso más simple, esto es el de las traslaciones de Clifford de  $T^3$ . En la figura 3a se dibujaron sobrepuestos los HSPM de  $T^3$  y de  $\mathcal{R}^3$ . En la figura 3b aparece la diferencia entre ambos histogramas. En este caso hemos considerado  $\sim 35$  objetos en el poliedro fundamental, lo que resulta en  $N \simeq 150$  objetos en  $\mathcal{B}_R$  debido a las isometrías de traslación.

En el HSPM de  $T^3$  (figura 3a) se observa que la señal topológica debido a las traslaciones de Clifford se pone de manifiesto a través de los picos agudos, los cuales aparecen exactamente en  $\ell = 1, \sqrt{2}, \sqrt{3}$  que corresponden a las traslaciones  $\lambda = 1, \sqrt{2}, \sqrt{3}$ .



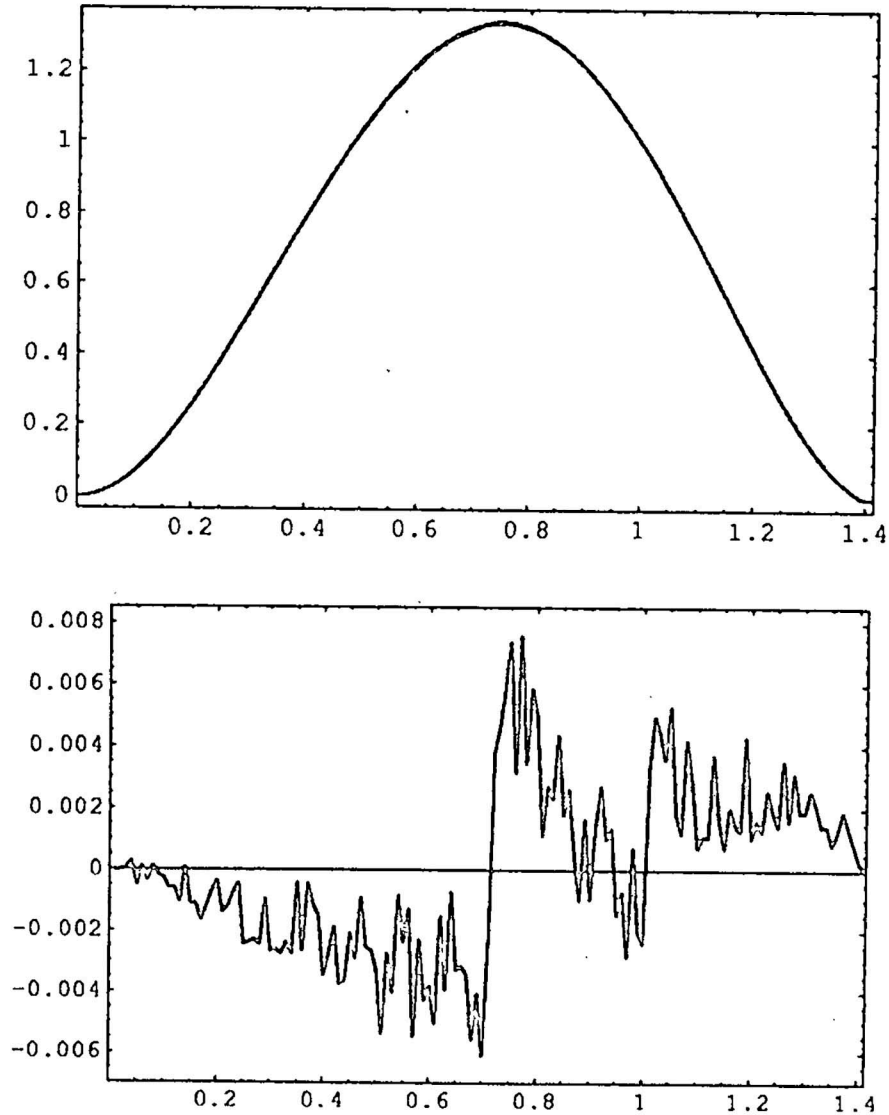
*Figura 3. (a) HSPM de  $T^3$  junto con el HSPM de  $\mathcal{R}^3$ . (b) Diferencia entre los histogramas de (a); se observan claramente los picos debido a las isometrías de traslación pura en  $\ell = 1, \sqrt{2}, \sqrt{3}$ .*

Ahora estudiaremos la isometría de *screw motion* del 3-espacio  $T_\pi$ . Si observamos el poliedro fundamental de este espacio, al lado derecho del tri-toro en la figura 1, notaremos que la cara inferior  $F$  es trasladada y rotada para identificarse con la cara superior. Pero simultáneamente ocurre que los otros dos pares de caras se identifican con traslaciones puras con  $\lambda = 1$  y  $\lambda = \sqrt{2}$ . En consecuencia, aparecen sendos picos agudos en  $\ell = 1$ ,  $\sqrt{2}$  y no parece ninguno en  $\ell = \sqrt{3}$ . Además hay una diferencia notoria entre el pico agudo en  $\ell = 1$  que aparece en el tri-toro comparado con el que aparece en la misma posición en  $T_\pi$ . En este último caso hay dos efectos sobrepuestos: la traslación pura de longitud  $\ell = 1$  debido a las caras  $G$  y  $J$  que se identifican por traslación pura, sumado al efecto de *screw motion* de la cara  $F$ . Esta diferencia se aprecia claramente en la figura 4b (comparar con la figura 3b).



**Figura 4.** (a) HSPM de  $T_\pi$  junto con el HSPM de  $\mathcal{R}^3$ . (b) Diferencia entre los histogramas de (a); se observan claramente las isometrías de *screw motion* en  $\ell = 1$  y la de traslación pura en  $\ell = \sqrt{2}$ .

Finalmente estudiaremos las isometrías no-traslacionales de  $\mathcal{G}_6$ , dadas en las ecuaciones (23)-(25). En este caso el radio de  $\mathcal{B}_R$  ha sido escogido a propósito para mostrar en el HSPM solo isometrías no-traslacionales, es decir  $R = \sqrt{2}/2$ . Los resultados se observan en la figura 5.



*Figura 5. (a) HSPM de  $\mathcal{G}_6$  junto con el HSPM de  $\mathcal{R}^3$ ; la diferencia entre ambas curvas no es evidente. (b) Diferencia entre los histogramas medios de  $\mathcal{G}_6$  y  $\mathcal{R}^3$ ; en este gráfico sí se observa nítidamente la diferencia entre ambos histogramas debido justamente a las isometrías no-traslacionales en  $\ell = \sqrt{2} \simeq 0,71$  y en  $\ell = 1$ .*

## Conclusiones

Podemos concluir lo siguiente:

- El método de Cristalografía Cósmica es un buen método ya que nos muestra que la distribución de objetos es característica del 3-espacio que los contiene. Esto es, en dos 3-espacios con propiedades globales diferentes los objetos se distribuyen de manera también diferente.

Aquí hemos visto, por medio de los histogramas, de que manera es posible conocer algunas de las propiedades (denominadas isometrías) del 3-espacio que contiene los objetos.

- Existe sin embargo un serio problema para la aplicabilidad de este método al universo: no existen catálogos astronómicos completos, es decir con información de las posiciones espaciales de todos los objetos cósmicos localizados entre nosotros y la Última Superficie de Dispersión. Esto se debe principalmente a los siguientes factores,
  - a la evolución del universo, ya que los objetos cósmicos ahora observados no tuvieron siempre la estructura que hoy en día los caracteriza. Y como observar un objeto muy lejano significa observar hacia tiempos muy remotos –debido a la finitud de la velocidad de la luz– entonces resulta difícil saber cuál objeto cósmico hoy catalogado es imagen de otro;
  - los objetos cósmicos más lejanos observados –i.e. los cuasares, y solo unos pocos– se localizan aproximadamente a la mitad del radio del universo observable;
  - el plano de nuestra galaxia –la vía láctea– oscurece todo objeto ubicado detrás de él en el rango del espectro visible y de microondas;
  - no es posible medir la distancia de un objeto a nosotros, solo su posición angular en la esfera celeste y su corrimiento al rojo (redshift); el cálculo de distancias depende del modelo cosmológico que se adopte, el cual a su vez depende de los denominados parámetros cosmológicos, aún no determinados con suficiente precisión.
- La estrategia de promediar histogramas parece una buena opción para minimizar el ruido estadístico y evidenciar así las señales topológicas en los histogramas.
- Dado que no son viables los catálogos observacionales con información tri-dimensional de cada objeto cósmico (i.e. sus 3 coordenadas espaciales ( $r$ ,  $\theta$ ,  $\phi$ )), en la práctica no se pueden construir los HSP ni podemos pretender

encontrar correlaciones de distancia entre objetos. Debemos en cambio buscar correlaciones angulares entre los objetos catalogados con solo sus coordenadas angulares ( $\theta$ ,  $\varphi$ ) de ubicación en la esfera celeste.

Los catálogos más convenientes para este análisis son los mapas de temperatura de la Radiación Cósmica de Fondo [26, 27].

## BIBLIOGRAFÍA

- [1] Schwarzschild, K. 1900, *Vierteljahrsschrift der Ast. Ges.*, 35, 337 (english translation: 1998, *CQG*, 15, 2539).
- [2] de Sitter, W. 1917, *Proc. R. Acad. Amsterdam*, 20, 29
- [3] Ellis, G.F.R. 1971, *GRG*, 2, 7
- [4] Weeks, J.R. 1985, *The Shape of Space, Pure and Applied Math.*, Vol. 26 (New York: Marcel Dekker Inc.)
- [5] Ellis, G.F.R., & Schreiber, G. 1986, *Phys. Lett. A*, 115, 97.
- [6] de Oliveira-Costa, A., & Smoot, G.F. 1995, *Ap J*, 448, 477.
- [7] *Proceedings of Cleveland Workshop on Topology and Cosmology*, Oct. 1997 (1998, *CQG*, 15, 2529-2731).
- [8] Friedmann, A. 1922, *Z. Phys.*, 10, 377.
- [9] Lemaître, G. 1927, *Ann. Soc. Sci. Bruxelles*, 47, 29.
- [10] Misner, C., Thorne, K., & Wheeler, J.A., (Freeman & sons 1973, San Francisco).
- [11] Wolf, J.A. 1984, *Spaces of Constant Curvature* (Delaware: Publish or Perish Inc.).
- [12] Norberg, P., & Cole, S. 2000, *Science News*, 158, 104.
- [13] Lachièze-Rey, M., & Luminet, J. P. 1995, *Phys. Rep.*, 254, 135.
- [14] Lehoucq, R., Lachièze-Rey, M., & Luminet, J. P. 1996, *A & A*, 313, 339.
- [15] Lehoucq, R., Luminet, J.P., & Uzan, J.P. 1999, *Astr. & Astrophys.*, 344, 735.
- [16] Lehoucq, R., Uzan, J.P., & Luminet, J.P. 2000, preprint (astro-ph/0005515).
- [17] Fagundes, H.V., & Gausmann, E. 1998, preprint (astro-ph/9811368).
- [18] Fagundes, H.V., & Gausmann, E. 1998, *Phys. Lett. A*, 238, 235.
- [19] Gomero, G.I., Rebouças, M.J., & Teixeira, A.F.F. 2000, *Phys. Lett. A*, 275, 355.
- [20] Gomero, G.I., Rebouças, M.J., & Teixeira, A.F.F. 2000, *IJMPD*, 9, 687.
- [21] Gomero, G.I., Rebouças, M.J., & Teixeira, A.F.F. 2001, *CQG*, 18, 1885.
- [22] Gomero, G.I., Teixeira, A.F.F., Rebouças, M.J., & Bernui, A. 2002, *IJMPD*, 11, 869.
- [23] Rebouças, M.J. 2000, *IJMPD*, 9, 561.
- [24] Bernui, A., & Teixeira, A.F.F. 1999, preprint (astro-ph/9904180)
- [25] Bernui A., Gomero, G.I., Rebouças, M.J., & Teixeira, A.F.F. 1998, *Phys. Rev. D.*, 57, 4699.
- [26] MAP: <http://map.gsfc.nasa.gov/>
- [27] PLANCK: <http://astro.estec.esa.nl/Planck/>



# El sol visto desde la tierra

---

*Daniel Reyes Castillo\**

## *RESUMEN*

El presente trabajo estudia el movimiento del Sol visto por un observador en Tierra. Comenzamos transformando la ecuación de movimiento de la Tierra respecto al Sol –deducida usando la mecánica Newtoniana– para obtener las ecuaciones de movimiento del Sol en la esfera celeste. De estas identificamos la ecuación de la órbita periódica (en realidad no tan periódica) del Sol y la ecuación del movimiento del plano de ésta. Analizando estas ecuaciones, hallamos la velocidad de la órbita, la cual luego asociamos a la duración de los días. También deducimos los máximos y mínimos de la velocidad de la órbita y de la posición del plano de esta órbita.

## *ABSTRACT*

In this work we study Suns motion seen by an observer on Earth. We transform the Earths motion equation refered to the Sun –deduced from Newtonian mechanics– to obtain the Suns motion equation for the celestial sphere. Then we identify the Suns periodic equation (not so periodic actually) and its plane motion equation. In the analysis of these, we find the orbits velocity, which is then associated to days duration. We also deduce the maximums and minimums of the orbits velocity and of its plane position.

---

\* Universidad Nacional de Ingeniería - Facultad de Ciencias

\*danielreyes4@hotmail.com

## Introducción

La inclinación del eje terrestre influye en la periodicidad con la que deberíamos de ver al Sol (influyen también otras cosas más). Para saber como esto afecta la posición del Sol respecto a la Tierra, se necesitan conocimientos no muy avanzados de física y matemática. Necesitamos la trayectoria de la Tierra respecto del Sol, cambios de sistemas de referencia (por rotación y traslación), cambios de coordenadas (de rectangulares a esféricas), un poco de trigonometría y cálculo diferencial.

Hallaremos varias cosas:

- La posición del Sol, en la esfera celeste.
- El intervalo en el que se mueve el plano de rotación del Sol.
- La velocidad angular de rotación.
- Los lugares (en la trayectoria de la tierra respecto al Sol) donde la velocidad angular de rotación tiene sus valores extremos.

## La órbita Newtoniana terrestre respecto al Sol

El punto de partida es la representación clásica de la trayectoria terrestre respecto al Sol, es decir una elipse de excentricidad  $\epsilon$ , en un sistema de referencia  $O'$ .

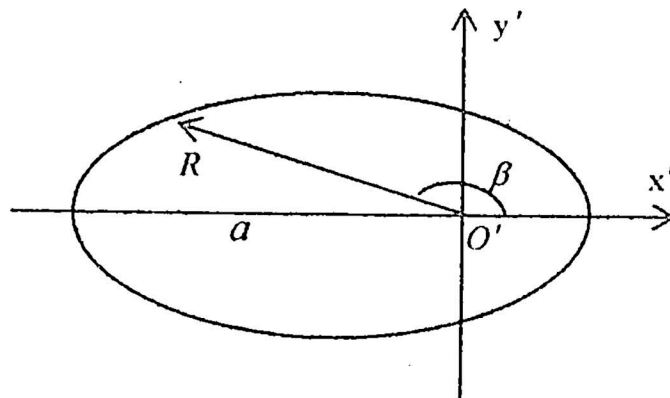


Figura 1

Del gráfico se observa que

$$\begin{aligned} x' &= R \cos \beta \\ y' &= R \sin \beta \\ z' &= 0. \end{aligned} \tag{1}$$

Vamos a tomar dos resultados muy conocidos de la gravitación clásica

$$R^2 \dot{\beta} = h = cte$$

$$R = \frac{a(1 - \varepsilon^2)}{1 + \varepsilon \cos \beta}$$

donde el punto denota derivada temporal. Combinando ambas, obtenemos un resultado que utilizaremos más adelante

$$\dot{\beta} = k(1 + \varepsilon \cos \beta)^2 \quad (2)$$

donde

$$k \equiv \frac{h}{a^2(1 - \varepsilon^2)^2}. \quad (3)$$

Primero necesitamos trasladar el origen de coordenadas que está en el Sol hacia la Tierra ( $O' \rightarrow O''$ ).

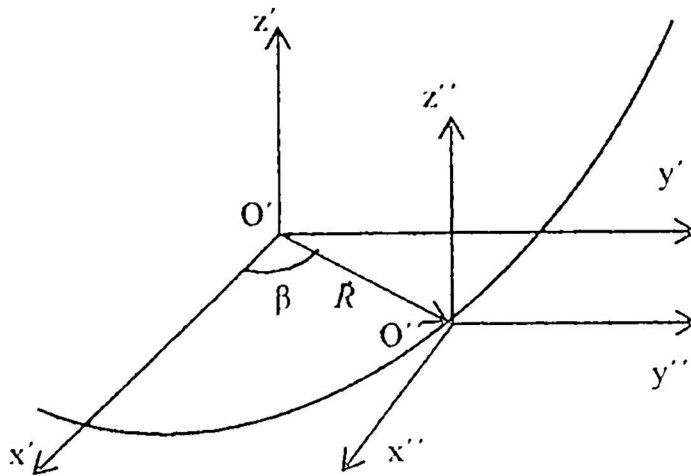


Figura 2

sabemos que para cambios por traslación

$$\vec{r}'' = \vec{r}' - \vec{R} \quad (4)$$

y como estamos trabajando con la posición del Sol  $\vec{r}' = \vec{0}$  entonces

$$\vec{r}'' = -\vec{R} \quad (5)$$

usando (5) en (1) obtenemos

$$\begin{aligned}x'' &= -R \cos \beta \\y'' &= -R \sen \beta \\z'' &= 0.\end{aligned}\tag{6}$$

Ahora hacemos una rotación por la inclinación del eje terrestre ( $O'' \rightarrow O'''$ )

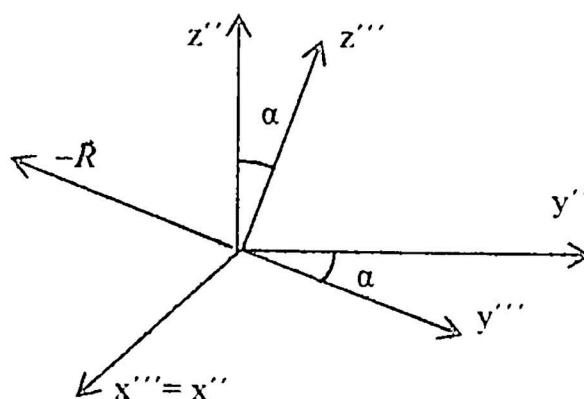


Figura 3

para hacer este cambio de sistema de referencia, usamos una matriz de rotación alrededor del eje  $x''$

$$\begin{pmatrix} x''' \\ y''' \\ z''' \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos \alpha & -\sen \alpha \\ 0 & \sen \alpha & \cos \alpha \end{pmatrix} \begin{pmatrix} x'' \\ y'' \\ z'' \end{pmatrix}\tag{7}$$

reemplazamos (6) en (7) y resulta

$$\begin{aligned}x''' &= -R \cos \beta \\y''' &= -R \sen \beta \cos \alpha \\z''' &= -R \sen \beta \sen \alpha.\end{aligned}\tag{8}$$

La segunda rotación va a ser alrededor del eje  $z'''$ , que es el eje de rotación terrestre; y, va a depender del tiempo, para que el nuevo SR gire junto con la Tierra ( $O''' \rightarrow O$ ).

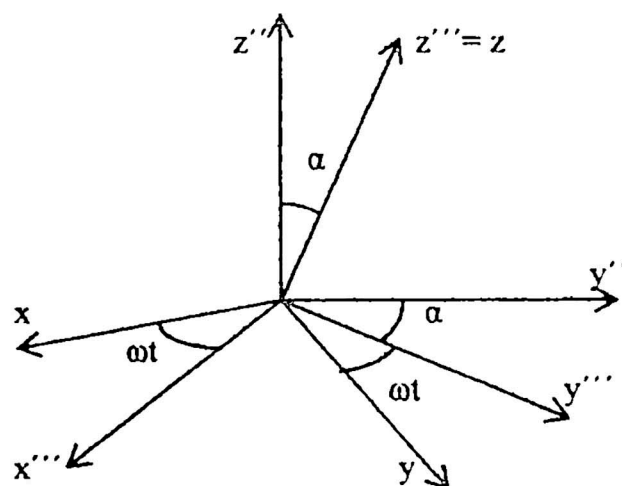


Figura 4

esto lo hacemos mediante otra matriz

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} \cos \omega t & \sin \omega t & 0 \\ -\sin \omega t & \cos \omega t & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x'' \\ y'' \\ z'' \end{pmatrix} \quad (9)$$

reemplazamos (8) en (9)

$$\begin{aligned} x &= -R \cos \beta \cos \omega t - R \sin \beta \cos \alpha \sin \omega t \\ y &= R \cos \beta \sin \omega t - R \cos \beta \cos \omega t \\ z &= -R \sin \beta \sin \alpha. \end{aligned} \quad (10)$$

Ahora que ya tenemos la posición del Sol en coordenadas rectangulares, debemos expresarla en coordenadas esféricas para obviar la coordenada radial, ya que sólo nos interesa la posición del Sol en la esfera celeste. Sabemos que

$$\operatorname{tg} \theta = \frac{\sqrt{x^2 + y^2}}{z} \quad (11)$$

$$\operatorname{tg} \phi = \frac{y}{x} \quad (12)$$

usando las ecuaciones (10) en (11) y (12)

$$\theta = - \operatorname{arc} \operatorname{tg} \left( \frac{\sqrt{1 - \operatorname{sen}^2 \alpha \operatorname{sen}^2 \beta}}{\operatorname{sen} \beta \operatorname{sen} \alpha} \right) \quad (13)$$

$$\phi = -\omega t - \frac{\pi}{2} + \operatorname{arc} \operatorname{tg} \left( \frac{\cos \beta}{\operatorname{sen} \beta \cos \alpha} \right) \quad (14)$$

considerando la siguiente figura y la forma de la ecuación (14), observamos que el ángulo  $\phi$  es el que describe la rotación del Sol respecto a la Tierra ya que depende de  $\omega$ . Por otro lado el ángulo  $\theta$ , ecuación (13), es el que nos indica la posición del “plano” en que rota el Sol respecto a la Tierra.

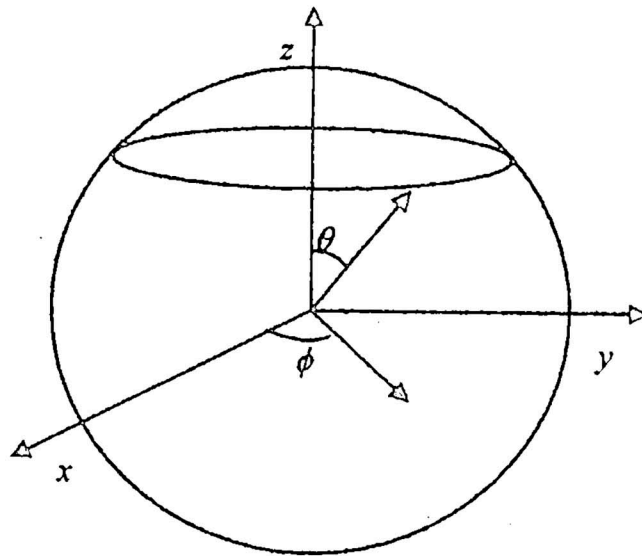


Figura 5

Para hallar la velocidad angular con la que vemos rotar al Sol, derivamos (14) con respecto al tiempo y operando

$$\dot{\phi} = -\omega + \frac{\dot{\beta} \cos \alpha}{1 - \operatorname{sen}^2 \alpha \operatorname{sen}^2 \beta} \quad (15)$$

usando (3) en (15)

$$\dot{\phi} = -\omega + \frac{k \cos \alpha (1 + \varepsilon \cos \beta)^2}{1 - \operatorname{sen}^2 \alpha \operatorname{sen}^2 \beta} \quad (16)$$

Ahora hallaremos los puntos donde esta velocidad angular toma sus valores extremos. Derivando la ecuación (16), igualándola a cero y simplificando

$$-\varepsilon \sin \beta (1 - \sin^2 \alpha \sin^2 \beta) + (1 + \varepsilon \cos \beta) \sin^2 \alpha \sin \beta \cos \beta = 0 \quad (17)$$

se observa que cuando  $\beta$  toma los valores de 0 (punto  $V_1$  de la figura 6) y  $\pi$  (punto  $V_3$ ) la ecuación se anula, entonces son soluciones de la ecuación (17). Para buscar otras soluciones asumimos  $\beta$  diferente de 0 y  $\pi$ , entonces la ecuación se puede reducir a

$$\cos \beta = \varepsilon \operatorname{ctg}^2 \alpha \quad (18)$$

como se sabe que  $\varepsilon = 0,017$  y  $\alpha = 23,5^\circ$  entonces

$$\cos \beta = 0,089 \quad (19)$$

de aquí encontramos otras dos soluciones para  $\beta$  :  $84,84^\circ$  ( $V_2$ ) y  $275,16^\circ$  ( $V_4$ ).

Por último para encontrar el intervalo en que se mueve el plano de rotación del Sol en la esfera celeste, podríamos derivar con respecto a  $\beta$  la ecuación (13) e igualar a cero, siguiendo el mismo procedimiento anterior; pero un poco de análisis nos facilitará las cosas. En la misma ecuación (13) el menor valor del ángulo  $\theta$  ( $\pi/2 - \alpha$ ) se da si  $\beta$  es  $\pi/2$  (punto  $P_1$  de la figura 6); y su mayor valor ( $\pi/2 + \alpha$ ), si  $\beta$  es  $3\pi/2$  (punto  $P_2$ ).

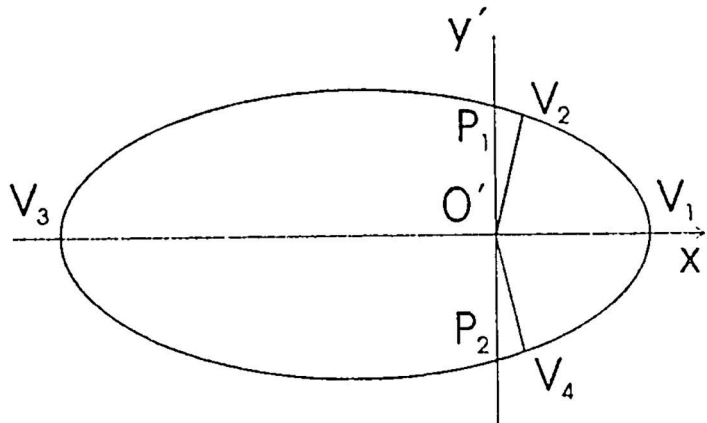


Figura 6

## Conclusiones

Hemos encontrado las ecuaciones que describen la posición del Sol en la esfera celeste (13) y (14). Como era de esperarse las ecuaciones de  $\theta$  y de  $\phi$  dependen de la posición de la Tierra respecto al Sol (debido al ángulo  $\beta$ ). De estas ecuaciones obtuvimos que  $\theta$ , que determina la posición del "plano" de rotación del Sol, pertenece al intervalo  $[\pi/2 - \alpha, \pi/2 + \alpha]$ , donde  $\alpha$  es la inclinación del eje terrestre. Se ha colocado la palabra plano entre comillas; porque en realidad la trayectoria del Sol en ningún momento pertenece a un plano. Se ha utilizado este recurso porque es más fácil visualizar que el Sol está rotando en un plano en movimiento. En realidad, la proyección de esta trayectoria en la esfera celeste es más como un helicoides, estando nosotros en el centro de ella.

Además hemos hallado la expresión de la velocidad angular de rotación del Sol, ecuación (16). Se observa que al esperado valor de  $-\omega$  se le suma un término correctivo que depende de la posición de la Tierra respecto al Sol. Por último los lugares donde esta velocidad angular toma sus valores extremos están especificados en la figura 6.

## BIBLIOGRAFÍA

- [1] *Feynmann / Leighton / Sands 1987, Física, Vol. 1 (Addison-Wesley).*



# Medidas de *SEEING* en el observatorio de Huancayo

---

Antonio Pereyra<sup>1,2</sup>, Nobar Baella<sup>1</sup>

## RESUMEN

Presentamos medidas preliminares para la determinación del *seeing* (calidad de cielo) para observaciones astronómicas en el observatorio de Huancayo. Medidas de perfiles estelares son utilizadas para cuantificar el *seeing* y es mostrada su dependencia con la masa de aire. Un valor promedio representativo de 2.71" es encontrado con un valor mínimo medido de 1.76".

Las medidas fueron realizadas usando una cámara CCD.

## ABSTRACT

*We present preliminary determinations of seeing for astronomical observations at the Huancayo's Observatory. Measurements of star profiles are used to quantify the seeing and is showed its dependence with airmasses. A representative mean value of 2.71" is found with a measured minimum value of 1.76". The measurements were made using a CCD camera.*

---

<sup>1</sup> Grupo Astronomía - Facultad de Ciencias, Universidad Nacional de Ingeniería, Lima, Perú  
[nobar\\_octa@yahoo.com](mailto:nobar_octa@yahoo.com)

<sup>2</sup> Instituto de Astronomia, Geofísica e Ciências Atmosféricas, Universidade de São Paulo, São Paulo, Brasil. [antonio@astro.iag.usp.br](mailto:antonio@astro.iag.usp.br) [antonio@astro.iag.usp.br](mailto:antonio@astro.iag.usp.br)

## Introducción

El conocimiento de los factores meteorológicos y climáticos, de los cuales depende la calidad de imagen de los objetos astronómicos, es imprescindible en la elección de un determinado lugar que tenga previsto la instalación de instrumentos ópticos astronómicos. En general, los factores a evaluar son el grado de turbulencia atmosférica (*seeing*), la transparencia de la atmósfera y el nivel de luminosidad del fondo de cielo nocturno.

El frente de onda proveniente de un objeto celeste se deforma fundamentalmente en las capas bajas de la atmósfera, cercanas al suelo (capas de inversión), bien sea por el calentamiento de ellas (por el día) o por el enfriamiento (por la noche). Esto causa que se forme un gradiente de temperatura que es el responsable de la variación espacial y temporal del índice de refracción, que ocasiona así la deformación del frente de onda. La calidad de imagen puede estar afectada por cambios de posición, de tamaño y de intensidad (centelleo). En las cumbres aisladas y mesetas elevadas, el espesor de esa capa de inversión de temperatura puede alcanzar varias decenas de metros, mientras que en los valles y depresiones varios cientos de metros. De ahí que estos últimos lugares no sean buenos para la instalación de telescopios. Además, el aire que baja por las laderas de las elevaciones hacia valles y depresiones causa una turbulencia adicional que es perjudicial para la calidad de la imagen. Por la existencia de heterogeneidades de temperatura cercanas al suelo es que los telescopios se elevan algunos metros sobre éste. En las capas altas de la atmósfera (tropopausa) también ocurren fluctuaciones de la densidad y del índice de refracción del aire, pero los gradientes son menores que en las capas cercanas al suelo, por lo que no causan el cambio de tamaño de las imágenes, sino solo el centelleo y el cambio de posición (vibraciones) de las estrellas.

En el foco de un gran telescopio los efectos de turbulencia atmosférica forman rápidamente imágenes *speckle*. Una imagen de largo tiempo de integración es el resultado de la suma de un gran número de imágenes *speckle* aleatorias. Esto resulta en que un perfil estelar (caracterizado por su ancho total a media altura, *full width at half maximum*, *FWHM*), y que cuantifica al *seeing*, presente *FWHM* típicos de 0.5 a 2 segundos de arco para un buen sitio de observación.

El modelo patrón para *seeing* astronómico, desarrollado por Tatarski (1961) [1] y Fried (1965)[2], está basado en el trabajo de Kolmogorov (1941)[3] sobre turbulencia atmosférica. El análisis ha sido revisado en detalle por Roddier (1981)[4]. El resultado crucial es que la propagación de la turbulencia en el modelo de

Kolmogorov está dado por la función de estructura  $D_p(r)$  de las perturbaciones de la fase del frente de onda  $p(r)$  y que depende de la escala de separación  $r$  según.

$$D_p(r) = \langle [p(r') - p(r'-r)]^2 \rangle = 6.88 (|r|/r_0)^{5/3}$$

donde la escala de longitud  $r_0$  es conocido como el **parámetro de Fried**, y es una medida de la intensidad de las distorsiones del *seeing*. Para esta función de estructura, el *seeing* limita las *FWHM* en integraciones largas y para un telescopio con diámetro mucho mayor que  $r_0$ , según:

$$FWHM = 0.98 \lambda / r_0(z) \quad (1)$$

donde  $\lambda$  es la longitud de onda de la observación, y  $r_0(z)$  indica la dependencia de  $r_0$  con la distancia cenital,  $z$ . Es importante notar que la *FWHM* observada será igual a la predicha por la ecuación 1 solamente si no hay contribución al ensanchamiento de la imagen debido a otras fuentes tales como la desfocalización del telescopio o errores en el acompañamiento sideral. Para comparar medidas de *seeing* a diferentes  $z$ , debemos corregir el *FWHM* observado considerando que:

$$r_0(\text{cenit}) = r_0(z) \sec(z)^{0.6}$$

Con lo cual de la ecuación (1) tenemos:

$$\begin{aligned} FWHM &= 0.98 \lambda \sec(z)^{0.6} / r_0(\text{cenit}) \\ &\propto \sec(z)^{0.6} \end{aligned} \quad (2)$$

o bien,

$$FWHM_{\text{corregido}} = FWHM_{\text{observado}} / \sec(z)^{0.6} \quad (3)$$

Para  $z < 60^\circ$ , podemos usar la aproximación plano-paralela de la atmósfera, donde la masa de aire  $X$  queda definida así:

$$X = \sec(z) \quad (4)$$

siendo,

$$\sec z = (\sin \phi \sin \delta + \cos \phi \cos \delta \cos H)^{-1} \quad (5)$$

y donde  $\phi$  es la latitud del observador,  $\delta$  es la declinación del objeto y  $H$  es el ángulo horario del objeto.

## Datos observacionales

Las medidas fueron realizadas en el Observatorio de Huancayo ( $12^{\circ}02'12''$  latitud sur;  $75^{\circ}19'07''$  longitud oeste; 3300 msnm), localidad de Huayao, Huancayo, durante la noche del 17/18 de Septiembre de 2001. Fue utilizado un telescopio Schmidt-Cassegrain de 8" de apertura (Celestron Celestar 8), de razón focal  $f/10$ , instalado sobre una montura ecuatorial tipo horquilla dentro de la cúpula automática del observatorio. Una cámara CCD (SBIG ST-7E)<sup>3</sup> fue adaptada al foco del telescopio donde son formadas las imágenes estelares. La cámara posee un chip Kodak KAF-0401E (765 x 510 píxeles) con un tamaño de píxel de  $9\mu \times 9\mu$  y que para la distancia focal del telescopio ( $F = 2032$  mm) proporciona una escala de placa de  $0.91''/\text{píxel}$ . El campo total de cielo cubierto por el CCD es de  $12' \times 8'$ . El sistema de enfriamiento termoeléctrico de la cámara trabaja a  $25^{\circ}$  por debajo de la temperatura ambiente, la cual fue de aprox.  $10^{\circ}$  C al comienzo de la noche.

Las medidas fueron realizadas usando los filtros  $U$ ,  $B$  y  $V$  en las bandas ultravioleta, azul, y visual, respectivamente, del sistema fotométrico Johnson-Morgan. Los filtros fueron instalados en una rueda automática (CFW8) que permite un rápido intercambio de filtros durante las observaciones. La cámara CCD como la rueda de filtros son controlados a través del *software* de control CCDOPS (ver. 4.12-NT) instalado en una computadora portátil y donde la conexión es hecha usando el puerto paralelo. Las imágenes digitales generadas son almacenadas convenientemente para su posterior procesamiento.

Se presentaron algunos inconvenientes en la observación de estrellas en masas de aire mayores. Por un lado el área efectiva disponible de la abertura de la cúpula limita la observación a alturas pequeñas sobre el horizonte y, por otro, el poco espacio disponible en la montura tipo horquilla para la colocación de la cámara CCD más la rueda de filtros también limita el mismo tipo de observación.

El sistema telescopio, cámara y rueda de filtros fue contrapesado convenientemente para permitir una fácil localización de las estrellas a estudiar. La localización de los objetos estelares fue hecha manualmente utilizando un buscador de 7 x 50 (pequeño telescopio de 7 aumentos y 5 cm de apertura) y por comparación con cartas estelares electrónicas. Imágenes de calibración de *flat field* en cada filtro y *dark current* fueron tomadas al comienzo y al final de la noche.

<sup>3</sup> <http://www.sbig.com> (Santa Barbara Instrument Group, Inc.)

Fueron observadas un total de 7 estrellas de tipo A0 (o aproximado) en varias masas de aire. En cada filtro (*U*, *B* y *V*), fueron tomadas imágenes individuales para cada estrella. La Tabla 1 lista las estrellas observadas. En la columna (1) está el número de observaciones realizadas. Todas las estrellas fueron observadas en una única secuencia de filtros a una determinada masa de aire, excepto las estrellas SAO 144150 (observaciones 3 y 4) y SAO 146044 (observaciones 6 y 7) que fueron observadas en dos masas de aire diferentes. En la columna (2) está el nombre de la estrella respectiva según el catálogo SAO<sup>4</sup>; en las columnas (3) y (4) están indicadas la coordenadas ecuatoriales (2000) para cada objeto; la columna (5) indica el tipo espectral de cada estrella según la base de datos astronómica *SIMBAD*<sup>5</sup>; la columna (6) indica el filtro utilizado; las columnas (7) y (8) indican la fecha y hora local de la observación; la columna (9) muestra la masa de aire (*X*) calculada, para la posición del objeto al momento de ser observado; la columna (10) muestra el valor de *seeing* observado ( $FWHM_{obs}$ ) medido para cada objeto en segundos de arco; y, finalmente, la columna (11) muestra el *seeing* corregido por la masa de aire ( $FWHM_{cor}$ ) según la ecuación 3.

Los tiempos de integración (TI) para todas las estrellas listadas en la Tabla 1 fueron de 1 segundo en los tres filtros. Las estrellas escogidas eran más brillantes que 6ta. magnitud para que integraciones rápidas puedan realizarse. Problemas en el motor de acompañamiento sideral y en el alineamiento polar actual no permiten hacer integraciones largas para objetos a distancias cenitales grandes.

---

<sup>4</sup> *Smithsonian Astrophysical Observatory Star Catalog*

<sup>5</sup> <http://simbad.u-strasbg.fr/Simbad>

*Tabla 1. Lista de estrellas observadas.*

observación (1)	estrella (2)	A. R. <sub>2000</sub> (3)	Dec. <sub>2000</sub> (4)	T. E. (5)	filtro (6)	día (7)	hora (8)	$X$ (9)	$fwhm_{obs}$ (10)	$fwhm_{cor}$ (11)
1	SAO 110543	02 28 09.54	+08 27 36.2	B9III	$U$	09/18/2001	01:53:28	1.0226	4.16	3.95
					$B$	09/18/2001	01:51:54	1.0240	2.88	2.74
					$V$	09/18/2001	01:48:43	1.0271	2.85	2.70
2	SAO 122754	17 47 53.56	+02 42 26.2	A0V	$U$	09/17/2001	22:02:34	2.0302	5.06	3.23
					$B$	09/17/2001	22:00:35	2.0004	3.95	2.54
					$V$	09/17/2001	21:58:26	1.9693	3.78	2.46
3	SAO 144150	20 11 18.29	-00 49 17.3	B9.5III	$U$	09/17/2001	21:32:55	1.0736	4.10	3.95
					$B$	09/17/2001	21:37:16	1.0802	3.91	3.75
					$V$	09/17/2001	21:39:10	1.0832	4.06	3.88
4					$U$	09/18/2001	00:13:49	1.9111	3.33	2.27
					$B$	09/18/2001	00:12:13	1.8903	3.39	2.33
					$V$	09/18/2001	00:09:29	1.8561	3.29	2.29
5	SAO 144810	20 47 40.55	-09 29 44.8	A1V	$U$	09/17/2001	21:18:36	1.0783	2.83	2.82
					$B$	09/17/2001	21:17:00	1.0778	4.50	4.49
					$V$	09/17/2001	21:15:24	1.0772	3.29	3.29
6	SAO 146044	22 21 39.38	-01 23 14.4	A0V	$U$	09/17/2001	22:40:19	1.0284	3.69	3.65
					$B$	09/17/2001	22:37:58	1.0282	3.90	3.86
					$V$	09/17/2001	22:34:24	1.0280	3.34	3.31
7					$U$	09/18/2001	00:33:49	1.1870	3.57	3.24
					$B$	09/18/2001	00:32:22	1.1826	2.64	2.40
					$V$	09/18/2001	00:30:45	1.1779	2.71	2.48
8	SAO 55397	02 15 56.28	+33 21 32.0	AOV	$U$	09/18/2001	01:18:21	1.1186	3.79	2.97
					$B$	09/18/2001	01:16:48	1.1207	2.53	1.98
					$V$	09/18/2001	01:14:40	1.1236	2.26	1.76
9	SAO 55427	02 17 18.87	+33 50 49.9	A1Vnn	$U$	09/18/2001	01:08:50	1.1382	4.15	3.19
					$B$	09/18/2001	01:06:59	1.1411	3.07	2.36
					$V$	09/18/2001	01:05:13	1.1441	2.96	2.27

## Reducción y análisis de datos

El proceso de reducción y análisis de las imágenes digitales fue realizado en el sistema operativo Linux RedHat 7.2, usando el ambiente *IRAF*<sup>6</sup> [4,5]. Después de la corrección rutinaria de *darks* y *flats* sobre las imágenes, se procedió al cálculo de las masas de aire para cada imagen.

Para el correcto cálculo de las masas de aire, es necesario actualizar los valores correctos de la fecha de observación, hora local (*local time*, LC), tiempo sidereal (*sideral time*, ST) y coordenadas ecuatoriales con su época respectiva del objeto observado. Con estos valores y conociendo las coordenadas geográficas del sitio de observación, la masa de aire puede ser calculada, por ejemplo, utilizando la rutina *setairmass*<sup>7</sup> de *IRAF*. Los valores de la columna (9) de la Tabla 1 fueron calculados de esta manera. Esta rutina está basada en la ecuación 5.

El siguiente paso es analizar los perfiles estelares para determinar el *fwhm* en cada imagen de cada estrella. La Figura 1 muestra una imagen estelar típica (Fig. 1, arriba) producida por el CCD. El perfil estelar (Fig. 1, abajo) es construido graficando los valores de intensidad de los píxeles alrededor del centroide (que define el pico de intensidad) en función de la distancia radial. Los puntos representan los valores de los píxeles y la curva es el mejor ajuste de una gaussiana a los datos que permite construir la *FWHM*. Para esto se ha usado la rutina *daoedit*<sup>8</sup> de *IRAF*. El *seeing* observado medido de esta manera y el valor corregido (según ecuación 3) están indicados en las columnas 10 y 11 de la Tabla 1. Los histogramas de la Figura 2 representan la distribución del *seeing* medida en cada uno de los filtros. Los valores mínimos, máximos y promedio en cada filtro están mostrados en la Tabla 2. Es notorio que los menores valores son encontrados en el filtro V. Esto es consistente con el hecho que la focalización fue hecha en este filtro y por consiguiente las imágenes deben estar ligeramente desenfocadas en los filtros U y B, lo que se traduce en un *seeing* mayor en estos filtros.

<sup>6</sup> *IRAF* (*Image Reduction and Analysis Facility*), es un sistema de software de propósito general para la reducción y análisis de datos astronómicos. *IRAF* es escrito por el *IRAF programming group* (<http://iraf.noao.edu>) en NOAO (*National Optical Astronomy Observatories*) en Tucson, Arizona. NOAO es operado por AURA (*Association of Universities for Research in Astronomy, Inc.*) bajo acuerdo cooperativo con la NSF (*National Science Foundation*).

<sup>7</sup> *setairmass* es rutina de [noao.astutil](http://noao.astutil)

<sup>8</sup> *daoedit* es rutina de [digiphot.daophot](http://digiphot.daophot)



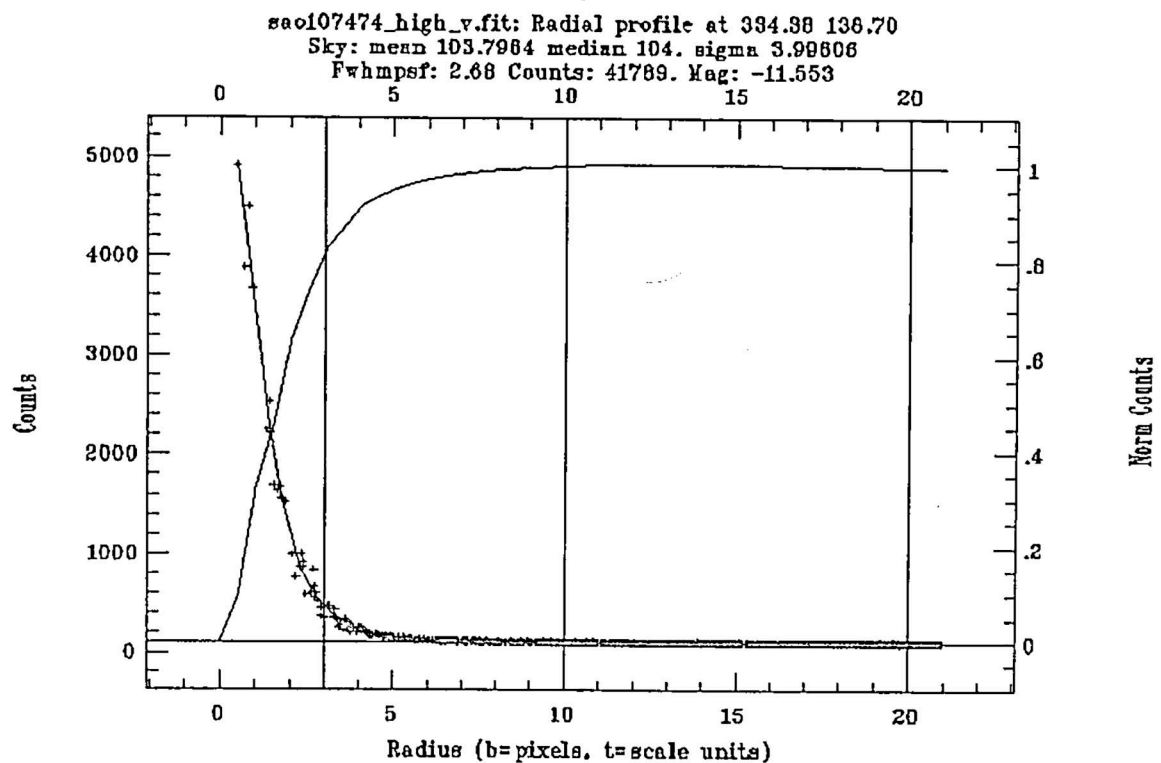
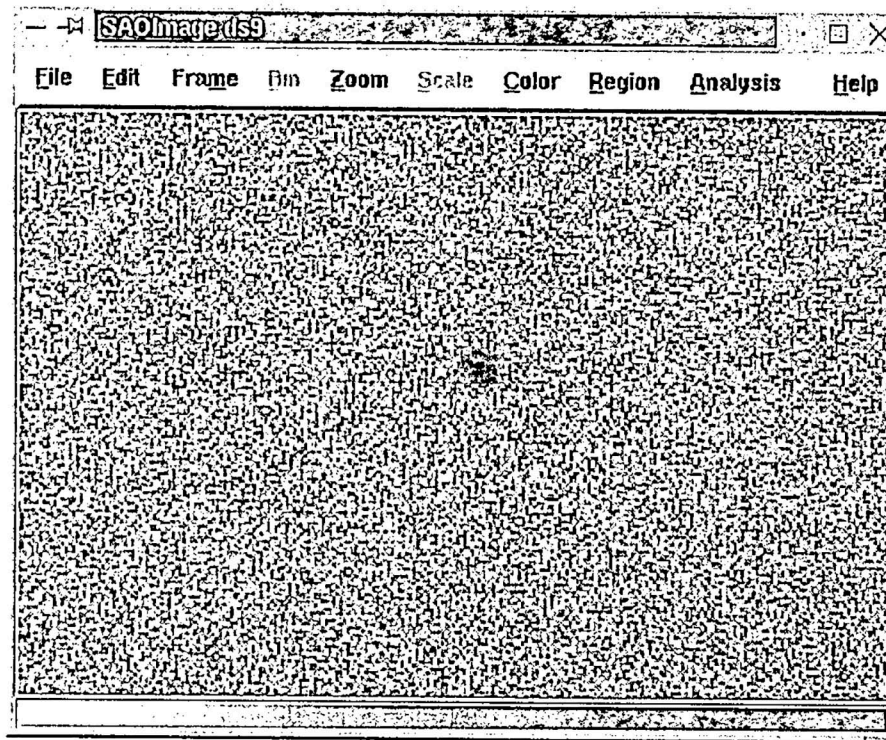
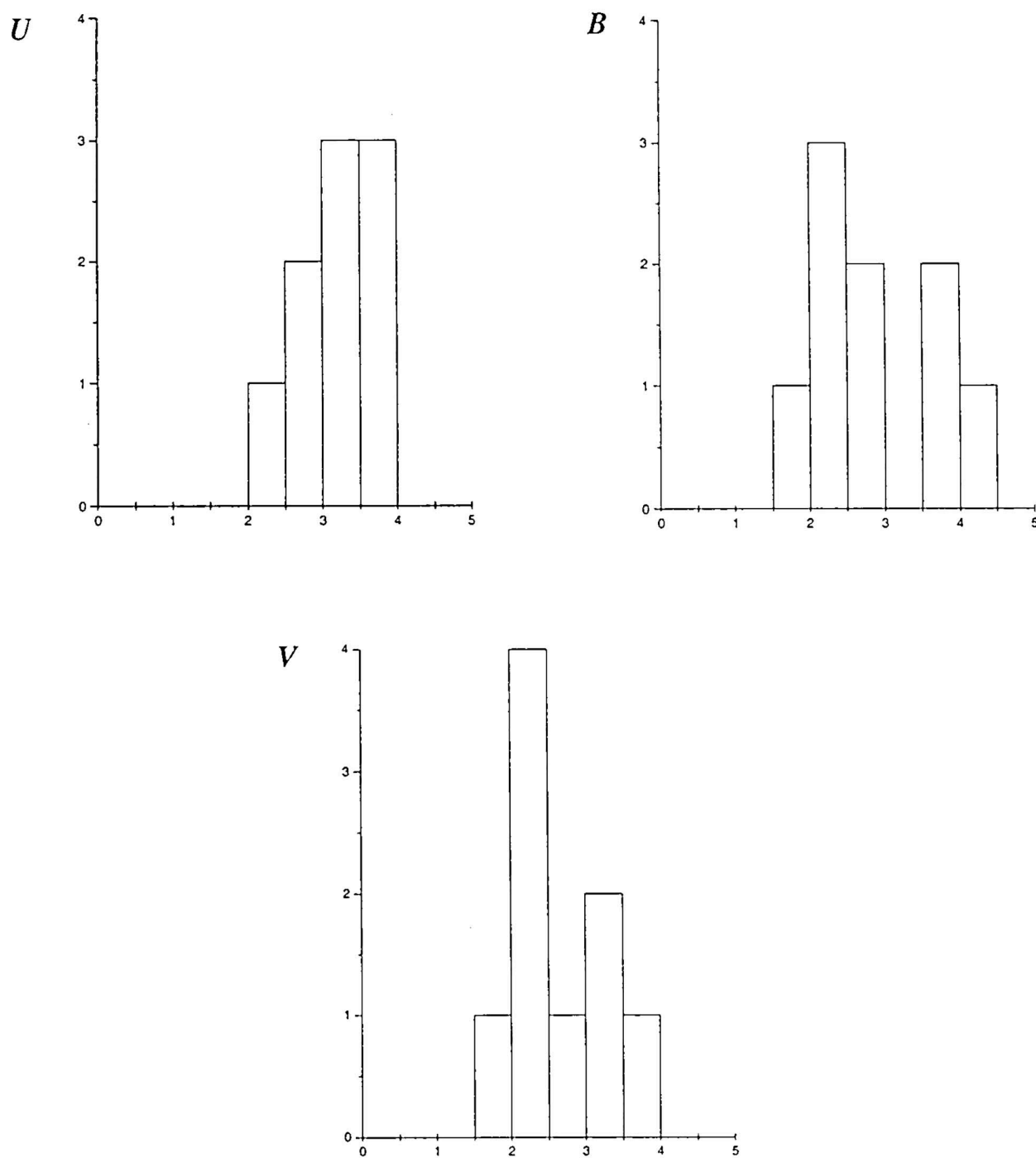


Figura 1. (arriba) Ejemplo de una imagen estelar sobre el CCD;  
 (abajo) Perfil radial para la imagen estelar y ajuste de gaussiana  
 para obtener la FWHM.



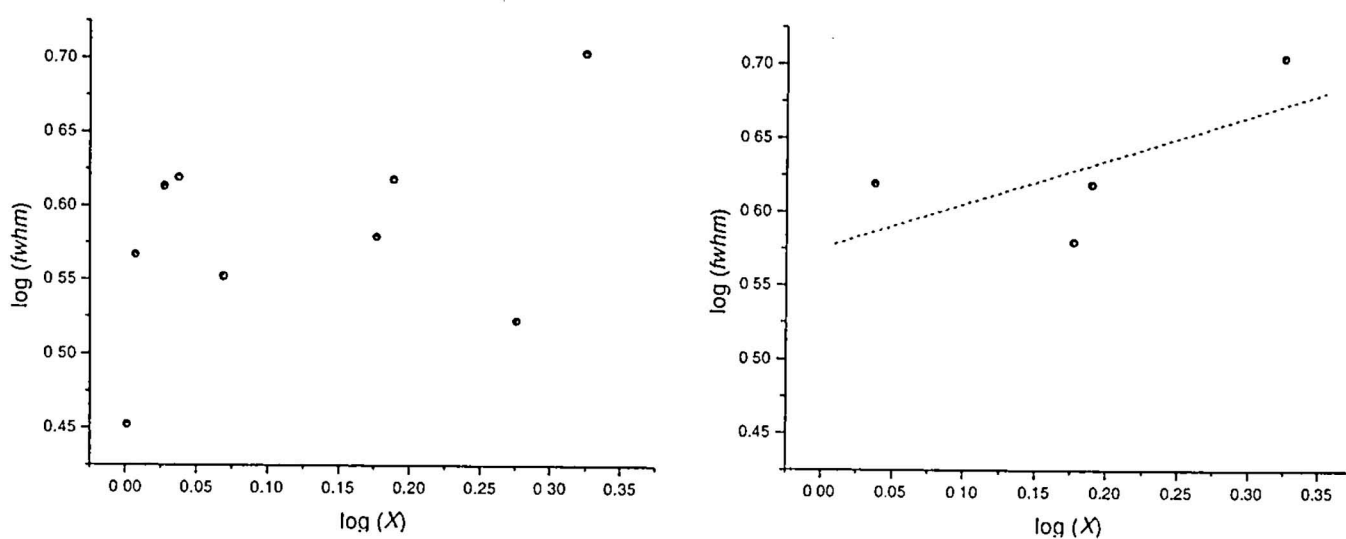


*Figura 2. Histogramas de los valores de seeing (corregidos por la masa de aire) para las medidas en los tres filtros: U, B y V. La abcisa está en segundos de arco.*

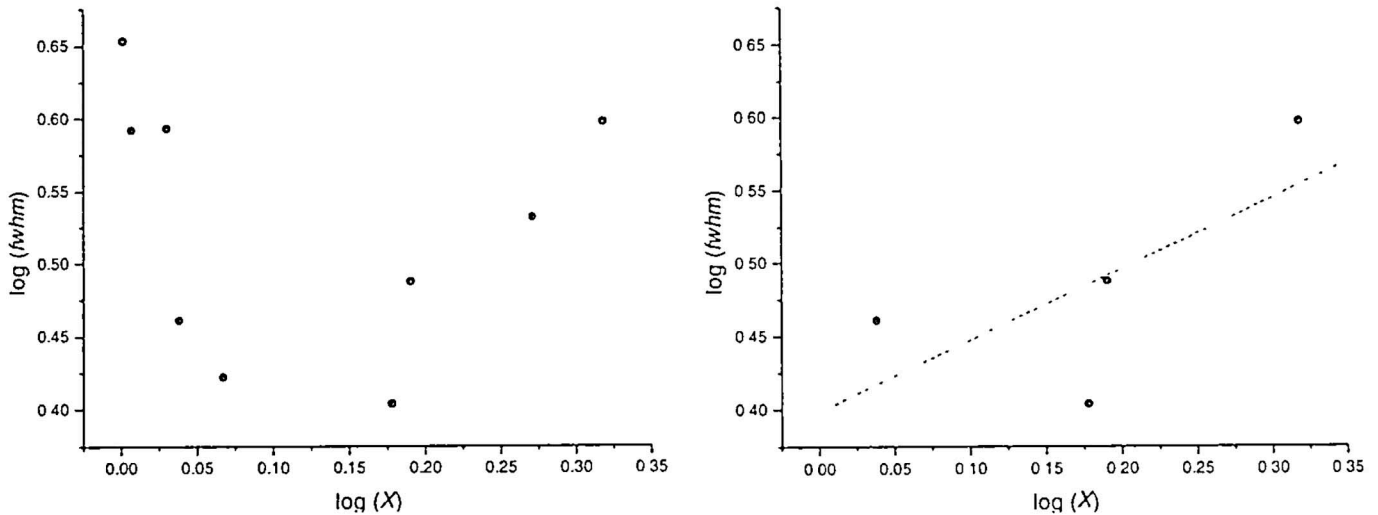
*Tabla 2. Estadísticas de los valores de seeing corregidos (por la masa de aire) para cada filtro. Todos los valores están en segundos de arco.*

	mínimo	máximo	promedio
<i>U</i>	2.27	3.95	3.25
<i>B</i>	1.98	4.49	2.94
<i>V</i>	1.76	3.88	2.71

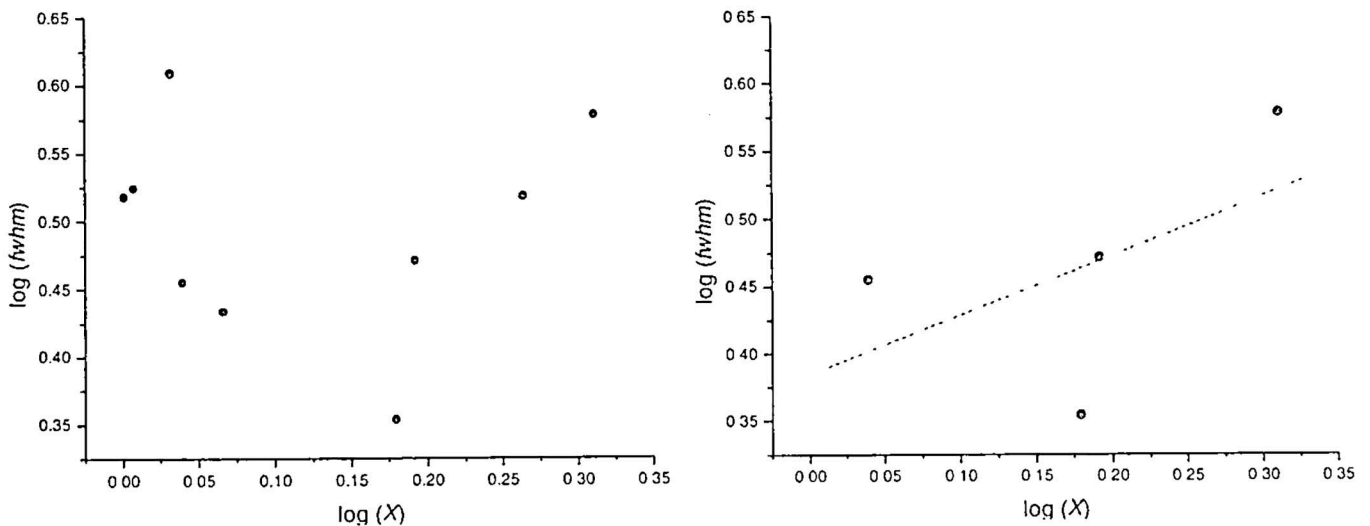
Para verificar la dependencia del *seeing* con la masa de aire (ecuación 3) graficamos el logaritmo de los valores de *seeing* observados -  $\log(fwhm_{obs})$  - vs.  $\log(X)$  y hacemos un ajuste lineal para determinar si la pendiente es consistente con el coeficiente 0.6 de la ecuación 3. Las Figuras 3, 4 y 5 (izquierda), grafican esta dependencia para todos los valores de la muestra. Una simple inspección nos indica que para masas de aire pequeñas el ajuste falla, pues valores altos de *seeing* son encontrados. Esto es más notorio en los filtros *B* y *V*. Para descartar el efecto de variabilidad aleatoria de la transparencia (efecto de paso de nubes pasajeras, por ejemplo), que pueda alterar las medidas en una secuencia típica de filtros para un objeto determinado, nos valemos de la información contenida en los tres filtros para cada objeto medido. Como quiera que la focalización fue realizada en el filtro *V* y es de esperar que, el filtro *B* presente menos desfocalización que el filtro *U*, procedemos a eliminar de la muestra aquellos objetos que presenten una patrón diferente a  $fwhm_V < fwhm_B < fwhm_U$ . Después de utilizar este criterio de filtrado



*Figura 3. (izquierda) Dependencia  $\log(fwhm)$  vs.  $\log(X)$  para todos los datos de la muestra tomados con el filtro *U*. (derecha) Ajuste lineal para datos filtrados. El valor calculado para la pendiente es  $0.29888 \pm 0.23695$ .*



**Figura 4.** (izquierda) Dependencia  $\log (fwhm)$  vs.  $\log (X)$  para todos los datos de la muestra tomados con el filtro B. (derecha) Ajuste lineal para datos filtrados. El valor calculado para la pendiente es  $0.48629 \pm 0.36464$ .



**Figura 5.** (izquierda) Dependencia  $\log (fwhm)$  vs.  $\log (X)$  para todos los datos de la muestra tomados con el filtro V. (derecha) Ajuste lineal para datos filtrados. El valor calculado para la pendiente es  $0.43954 \pm 0.49242$ .

apenas 4 objetos sobreviven en la muestra (observaciones 1, 2, 8 y 9) y 5 objetos son eliminados (observaciones 3, 4, 5, 6 y 7). Los ajuste lineales respectivos son mostrados en las Figuras 3, 4 y 5 (derecha). La pendiente obtenida en el filtro U ( $0.30 \pm 0.24$ ) está bastante lejana del valor esperado. Por otro lado, los ajustes en los filtros B y V ( $0.49 \pm 0.36$  y  $0.44 \pm 0.49$ , respectivamente), se encuentran en razonable concordancia con el valor esperado considerando los errores obtenidos. De cualquier manera, queda claro que la cantidad de datos utilizados es bastante reducida y se hace necesario más medidas para verificar los valores encontrados.

## Conclusiones

Una primera tentativa de medición de *seeing* en el Observatorio de Huancayo, usando una cámara CCD fue realizada. Los perfiles estelares caracterizados por la *FWHM* usados para las medidas de *seeing*, proporcionaron valores medios de 3.25", 2.94" y 2.71" medidos en los filtros *U*, *B* y *V* respectivamente. Las medidas en los filtros *U* y *B* sufrieron de ligera desfocalización por lo que las medidas en el filtro *V* son más representativas. Es interesante notar que un mínimo del orden de 1.76" fue medido en este filtro. Por otro lado, es bueno hacer notar que, a pesar de haber utilizado integraciones rápidas (1 seg.), es probable que problemas en el acompañamiento sideral de la montura utilizada así como en su alineamiento polar pueden haber condicionado a tener valores altos de *seeing*. Las condiciones atmosféricas durante la noche de observación estuvieron lejos de ser ideales, y variaciones en la transparencia del cielo (nubes pasajeras) estuvieron presentes. Una tentativa de caracterizar la dependencia del *seeing* con la masa de aire resultó en un pobre ajuste en el filtro *U* y apenas razonable en los filtros *B* y *V*, cuando son comparados a los valores previstos. Un mayor número de medidas bajo condiciones de cielo fotométrico se hacen necesarias para caracterizar mejor las medidas de *seeing* en el sitio.

## Agradecimientos

Los autores están especialmente agradecidos al personal del Observatorio de Huancayo por su colaboración para la realización de este trabajo. Este trabajo forma parte del proyecto "Observaciones Astronómicas con CCD" financiado por la Facultad de Ciencias – UNI e "Implementación del Laboratorio de Tratamiento de Imágenes Astronómicas CCD, FC-UNI", financiado por el CONCYTEC (C.S. No. 018-2002-CONCYTEC-OAJ). Este trabajo ha hecho uso de la base de datos *SIMBAD*, operada por el CDS (*Centre de Données astronomiques de Strasbourg*), Estrasburgo, Francia.

## REFERENCIAS

- [1] *Tatarski, V.I., 1961, Wavefront Propagation in a Turbulent Medium, Dover, New York.*
- [2] *Fried, D.L., 1965, J. Opt. Soc. Am., 55, 1427.*
- [3] *Kolmogorov, A.N., 1941, in Tikhomirov, V. M., ed, Selected works of A.N. Jolmogorov, Mathematics and its applications (Soviet series), Kluwer Academic press (1991).*
- [4] *Roddier, F., 1981, in Wolf E., ed, Progress in Optics 19, North Holland, Amsterdam.*

# El error inicial

---

H. G. Valqui \*

## RESUMEN

En 1935, Einstein y sus alumnos, plantearon una función de estado para el caso de dos partículas que habían cesado de interactuar entre sí. De dicho planteamiento nació la llamada paradoja EPR, según la cual entre dichas partículas existía un “enlace fantasmal”, tal que una medición sobre una de las partículas afectaba el estado de la otra partícula. De allí coligió Einstein que la mecánica cuántica era una teoría incompleta. En el presente artículo muestro la raíz del error cometido en el planteamiento de la función de onda.

## ABSTRACT

1935, Einstein, Podolski and Rosen presented a function representing the state of two particles which no longer interacted with each other. From that representation was born the so called EPR-paradox, according to which there is some sort of a ghost entanglement between both particles such that a particular measurement on one of the particles changes the state of the other one. Consequently, so Einstein, quantum mechanics is not a complete theory. In this article I show where is the root of the mistake made by constructing the state function.

Al considerar la ecuación de Schroedinger para dos partículas,  $C_1$  y  $C_2$ , que no interactúan entre sí <sup>(1)</sup>, la función de estado  $\Psi$  que describe el comportamiento simultáneo de ambas partículas puede ser escrita como el producto de las funciones de estado  $\chi$  y  $\eta$ , de cada una de las partículas, es decir,  $\Psi(x_1, x_2) = \chi(x_1) \eta(x_2)$ .

---

\* Facultad de Ciencias / UNI.

Pero ésta sería solamente una solución particular ; la solución general es la combinación lineal de todos los productos posibles, es decir,

$$\Psi(x_1, x_2) = \sum_{i,j} a_{ij} \chi_i(x_1) \eta_j(x_2) \quad [1]$$

donde, para cada situación física bien determinada los coeficientes  $a_{ij}$  quedarán fijados por las condiciones iniciales y las posibles condiciones de contorno del caso (en la presente situación no existen condiciones de contorno por tratarse de partículas libres).

Clasifiquemos los coeficientes  $a_{ij}$  en dos clases excluyentes:

- i) Para todo par de índices  $i, j$ , existen dos números  $b_i, c_j$ , tales que se puede escribir,  $a_{ij} = b_i c_j$ . En este caso diremos que los estados de  $C_1$  y de  $C_2$  son independientes.
- ii) Para todo par de índices  $i, j$ , NO existen dos números  $b_i, c_j$ , tales que se pueda escribir,  $a_{ij} = b_i c_j$ . En este caso diremos que los estados de  $C_1$  y de  $C_2$  se encuentran enlazados.

I) Estados independientes. En este caso podemos escribir

$$\Psi(x_1, x_2) = \sum_{i,j} b_i c_j \chi_i(x_1) \eta_j(x_2) = \psi_1(x_1) \psi_2(x_2), \text{ donde } \psi_1 = \sum_k b_k \chi_k, \psi_2 = \sum_k c_k \eta_k$$

Nótese que aquí podemos referirnos al estado de  $C_1$ , independientemente del estado de  $C_2$ , y recíprocamente.

Por otra parte, si, por ejemplo, sobre la partícula  $C_1$  se realizase alguna medición particular (no un proceso de medición) de un observable físico representado por el operador  $\Omega$ , cuyo resultado fuese el número  $\lambda$ , entonces, de acuerdo con el modelo cuántico, el estado de  $C_1$  sufriría un salto

$$\psi_1 \longrightarrow \phi, \quad \text{donde} \quad \Omega\phi = \lambda \phi$$

Como consecuencia de ello, también la función de estado común sufrirá el salto

$$\Psi = \psi_1 \psi_2 \longrightarrow \phi \psi_2 \text{ ó, si se prefiere, } \Psi(x_1, x_2) = \psi_1(x_1) \psi_2(x_2) \longrightarrow \phi(x_1) \psi_2(x_2)$$

donde el salto del estado de  $C_1$  no ha influido sobre el estado de  $C_2$ .

Por supuesto que si sobre la partícula  $C_2$  se realizase alguna medición particular, esto produciría el salto del estado de  $C_2$ , sin que sea afectado el estado de  $C_1$ .

De otro lado, se puede demostrar<sup>(2)</sup> que si en un instante dado, el estado común a dos partículas libres – como es el caso aquí considerado – se puede expresar como

el producto de los estados de las dos partículas en tal instante ; entonces, tal situación se mantendrá válida para todo instante posterior.

II) Estados Enlazados. En este caso propiamente no tiene sentido hablar del estado de la partícula  $C_1$  o de la partícula  $C_2$ ; solamente existe el estado común a las dos partículas. Por ejemplo, no puede hablarse del 'peso' que tiene el estado  $\chi_n$ , de  $C_1$ , en la descripción del estado común, porque dicho 'peso' debería estar expresado por un único coeficiente, y aquí estaría representado por todos los coeficientes  $\alpha_{nj}$ , es decir, para todo valor del índice  $j$ .

La situación de enlazamiento no causa ningún problema de entendimiento cuando las dos partículas se encuentran una muy cerca de la otra, como es el caso, por ejemplo, con la función de spin de los dos electrones de un átomo de helio ; aquí, si bien es cierto que (se asume) no existe un potencial (magnético) de interacción entre los dos electrones, en cambio los electrones no son libres, pues ambos están sometidos al potencial del núcleo.

El asunto es diferente cuando las partículas están alejadas, una de la otra. En este caso, no existiendo interacción entre ellas, se esperaría que cualquier acción sobre una de las partículas no afecte al estado de la otra partícula. Sin embargo, la expresión del estado en la forma [1] puede representar tanto estados independientes como estados enlazados. Aquí debemos tener presente que las fórmulas o las ecuaciones matemáticas pueden no representar situaciones físicas reales (como es el caso, por ejemplo, con el oscilador armónico para energías potenciales muy altas); siendo indispensable un cuidadoso análisis de interpretación por parte de los físicos (pues también existen casos, como es el del positrón, donde la solución matemática parece no tener sentido físico, cuando en realidad se trata sólo de limitaciones de interpretación, por falta de información experimental).

¿Cómo saber si una función de estado, común a dos (o más) partículas debe representar estados libres o deba representar estados enlazados? Aquí debemos tener presente que la función de estado [1] está representando el estado de dos partículas bien determinadas, donde se conocen los coeficientes  $\alpha_{ij}$ . Conociendo dichos coeficientes ya sabemos si se trata de estados independientes o estados enlazados.

¿Cómo se determinan los valores de los coeficientes  $\alpha_{ij}$ ?

Aquí se pone en evidencia que la representación [1] del estado común a dos partículas libres es deficiente; la función de estado deberá depender del tiempo,  $\Psi(x_1, x_2, t)$ , y deberá ser solución de la ecuación de Schroedinger

$$- (h^2/2m_1)\nabla_1^2\Psi - (h^2/2m_2)\nabla_2^2\Psi + \partial\Psi/\partial t = 0 \quad [2]$$

No existiendo potencial de interacción se puede ensayar la separación de variables  $\psi(x_1, x_2, t) = \chi(x_1)\eta(x_2)f(t)$ , que reemplazada en [2] genera las tres ecuaciones

$$-(\hbar^2/2m_1)\nabla_1^2 \chi = E_1 \chi, \quad -(\hbar^2/2m_2)\nabla_2^2 \eta = E_2 \eta, \quad (\hbar/i) \dot{f} = -(E_1 + E_2) f$$

donde  $E_1$  y  $E_2$  son constantes arbitrarias.

Para las ecuaciones anteriores podemos ensayar soluciones de las formas  $\chi_k(x_1) = \exp(i \mathbf{k} \cdot \mathbf{x}_1)$ ,  $\eta_q(x_2) = \exp(i \mathbf{q} \cdot \mathbf{x}_2)$ ,  $f(t) = \exp[-i \omega t]$

que introducidas correspondientemente en las ecuaciones dan las condiciones que deben satisfacer las constantes:  $E_1 = (\hbar k)^2/2m_1$ ,  $E_2 = (\hbar q)^2/2m_2$ ,  $E_1 + E_2 = \hbar \omega(\mathbf{k}, \mathbf{q})$ , donde los vectores  $\mathbf{k}$ ,  $\mathbf{q}$  son totalmente arbitrarios (y las otras constantes  $E_1$ ,  $E_2$ ,  $\omega$  se expresan en función de dichos vectores). Ahora, todos los miembros de la familia biparamétrica

$$\psi_{\mathbf{k} \mathbf{q}}(x_1, x_2, t) = \exp[i(\mathbf{k} \cdot \mathbf{x}_1 + \mathbf{q} \cdot \mathbf{x}_2 - i \omega(\mathbf{k}, \mathbf{q}) t)] \quad [3]$$

son soluciones particulares de la ecuación [2].

La solución del problema físico en consideración se obtendrá como una combinación lineal de dichas soluciones particulares:

$$\Psi(x_1, x_2, t) = (1/2\pi)^{3/2} \int d^3\mathbf{k} d^3\mathbf{q} \Phi(\mathbf{k}, \mathbf{q}) \exp[i(\mathbf{k} \cdot \mathbf{x}_1 + \mathbf{q} \cdot \mathbf{x}_2 - \omega(\mathbf{k}, \mathbf{q}) t)] \quad [4]$$

donde para conocer la función de estado es necesario conocer los coeficientes  $\Phi$ .

Del estado inicial,  $\Psi(x_1, x_2, 0) = (1/2\pi)^{3/2} \int d^3\mathbf{k} d^3\mathbf{q} \Phi(\mathbf{k}, \mathbf{q}) \exp[i(\mathbf{k} \cdot \mathbf{x}_1 + \mathbf{q} \cdot \mathbf{x}_2)]$  podemos apreciar que  $\Phi$  es la transformada de Fourier de  $\Psi_0$ , donde  $\Psi_0(x_1, x_2) \equiv \Psi(x_1, x_2, 0)$ ; es decir,

$$\Phi(\mathbf{k}, \mathbf{q}) = (1/2\pi)^{3/2} \int d^3\mathbf{x} d^3\mathbf{y} \Psi(x_1, x_2, 0) \exp[-i(\mathbf{k} \cdot \mathbf{x} + \mathbf{q} \cdot \mathbf{y})]$$

Es decir, para determinar la función de estado (y con ello saber si se trata de un caso de estados independientes o de estados enlazados) es necesario conocer el estado inicial común  $\Psi(x_1, x_2, 0)$ .

Pero también existe la alternativa experimental para saber si las dos partículas se encuentran en estados enlazados: actuar sobre una de las partículas y verificar si la otra partícula resulta afectada. Pero estas acciones sobre objetos microscópicos son indirectas y los resultados deben ser cuidadosamente interpretados a la luz de un conjunto de asunciones no necesariamente transparentes. Este es un tema que trataré en una próxima ocasión. Ahora paso al tema del título.

En 1935<sup>(3)</sup>, para representar la función de estado de dos partículas libres, Einstein procedió como sigue:



Sean  $u_1(x_1)$ ,  $u_2(x_1)$ ,  $u_3(x_1)$ , ... los estados propios de un observable A de un sistema  $C_1$ , donde  $x_1$  es la variable que describe el primer sistema. La función de estado  $\psi$ , considerada como una función de  $x_1$  puede ser expresada así

$$\psi(x_1, x_2) = \sum_k \phi_k(x_2) u_k(x_1)$$

donde  $x_2$  es la variable que describe el segundo sistema. Aquí los  $\phi_k(x_2)$  deben ser considerados simplemente como los coeficientes de la expansión de  $\psi$  en una serie de funciones ortogonales  $u_k(x_1)$ .

Es decir, Einstein usó una función de estado que es un caso muy particular de [1], donde  $\alpha_{ij} = \delta_{ij}$ , para todos los índices  $i, j$ , lo cual necesariamente implica un estado enlazado.

La función de estado planteada por Einstein tiene dos consecuencias significativas:

- i) Considerando que los estados de las dos partículas libres, alejadas una de la otra, deberían ser independientes y no enlazados, como resulta de la particular elección de los coeficientes, Einstein concluyó que la Mecánica Cuántica no podía ser una teoría completa, ya que no podía expresar la independencia de los estados de dos sistemas que no interactúan entre sí y están alejados el uno del otro.
- ii) Actualmente existen muchos físicos que sostienen la existencia de esos estados enlazados, y creen contar con razones experimentales que (interpretando adecuadamente los resultados experimentales) confirman la existencia de los mismos.

Finalmente debo mencionar que este error de Einstein tiene actualmente consecuencias de mucho peso en la concepción de la Teoría Cuántica y de sus aplicaciones. El error fue cometido en 1935, pero (según este autor) nunca ha sido considerado un error. En este sentido, cito como ejemplo, dos frases elocuentes de dos respetados físicos:

Clauser-Shymoni<sup>(4)</sup>: "In our opinion the reasoning of EPR is impeccable, once...."

A.Aspect<sup>(5)</sup>: "I found it extremely clear and completely convincing, but there was something special about this paper: it lead to two contradictory conclusions..."

#### NOTAS:


<sup>(1)</sup> Sin que tenga ninguna importancia que dichas partículas, antes del instante considerado como inicial, hayan estado interactuando entre sí.

<sup>(2)</sup> Esta demostración aparecerá en un siguiente artículo.

<sup>(3)</sup> Einstein, Podolski, Rosen, Can Quantum-Mechanical Description of Physical Reality Be Considered complete? , Physical Review, 47 , 1935.

<sup>(4)</sup> Clauser J. Shimomny A. , Bell's theorem: experimental tests and implications, Rep. Prog.Phys, Vol.41, 1978 (pág.1886)

<sup>(5)</sup> Aspect A., Testing Bell's Inequalities, College de France et Ecole Normal Supérieure (pág 416), con fecha posterior a Abril 1991.

 edición, diagramación e impresión por  
Editorial **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, febrero del 2003.*

## Í N D I C E

---

■	Raíces y álgebra de Lie semisimples diagramas de Dynkin <i>Aldo Arroyo Montero</i>	1
■	El método de Newton (amortiguado) para desigualdades variacionales. <i>Hermes Pantoja y William Echegaray</i>	19
■	Una prueba general de la buena definición del método lagrangeano aumentado. <i>Yna Consuelo Rezza Espinoza</i>	38
■	Películas de óxidos mixtos de cobre y tungsteno obtenidas por Sol-gel: caracterización estructural y evaluadas como sensor de vapor. <i>Angélica Damián Briones, Yovan Rodríguez Daga, José Solís, Walter Estrada</i>	58
■	Funciones de Green para un problema del electromagnetismo. <i>Carlos Enrique Valcárcel Flores</i>	70
■	Cristalografía Cósmica. <i>Armando Bernui</i>	77
■	El Sol visto desde la Tierra. <i>Daniel Reyes</i>	95
■	Medidas de SEEING en el observatorio de Huancayo. <i>Antonio Pereyra y Nobar Baella</i>	103
■	El error inicial. <i>H.G. Valqui</i>	115