UNIVERSIDAD DE LOS ANDES
FACULTAD DE CIENCIAS

NOTAS DE MATEMATICA

COLOQUIO MATEMATICO DE MERIDA

CONFERENCIAS
1992 - 1994

EDITORES

OSWALDO ARAUJO Y VICTOR PADRON

DEPARTAMENTO DE MATEMATICA
MERIDA - VENEZUELA

NOTAS DE MATEMATICAS

Nº 149

COLOQUIO MATEMATICO DE MERIDA

C O N F E R E N C I A S

1992 - 1994

EDITORES

OSWALDO ARAUJO Y VICTOR PADRON

UNIVERSIDAD DE LOS ANDES
FACULTAD DE CIENCIAS
DEPARTAMENTO DE MATEMATICA
MERIDA - VENEZUELA
1994

CONTENIDO

INTRODUCCION

El **COLOQUIO MATEMATICO DE MERIDA** es una actividad matemática auspiciada y promovida por el Posgrado del Departamento de Matemáticas de la Facultad de Ciencias de la Universidad de Los Andes (ULA), la Asociación Matemática Venezolana (AMV) y el Consejo de Desarrollo Científico, Humanístico y Tecnológico de la Universidad de Los Andes.

Este proyecto se inició en el año 1991 cuando los Profesores Gloria Sánchez, Carlos Uzcátegui, Jorge Vielma y Francisco Rivero, del Departamento de Matemáticas de la Facultad de Ciencias de la Universidad de Los Andes (ULA), tuvieron la idea de organizar un ciclo de conferencias de matemáticas dirigidas a un público no especializado. Se trataba de crear un espacio donde los diferentes miembros del Departamento hablaran de sus trabajos con el fin de generar una mayor comunicación entre ellos. En realidad, se intentaba revivir y darle organicidad a una actividad que el Departamento había iniciado en el pasado, pero que distintos factores, como excesivos compromisos docentes, la truncaron. Este ciclo de charlas fue llamado **COLOQUIO MATEMATICO ULA.**

Gracias a la existencia de un posgrado en matemática en distintas Instituciones de Educación Superior, visitan el país matemáticos de diferentes continentes, para dictar cursos o participar en eventos científicos. Su presencia, sumada a la de los matemáticos que trabajan en Venezuela, conforman un espectro de expositores potenciales de diversas áreas de la matemática. Esta circunstancia llevó al Capítulo Central de la AMV, a organizar el Coloquio Matemático de Caracas. Este evento consistía en programar conferencias con los matemáticos que se encontraban en Caracas y, en menor grado, con los que laboraban en el interior del país. El propósito era maximizar la inversión que una determinada Institución, esencialmente de la región capital, realizaba en la investigación y en docencia de posgrado.

La consideración del ejemplo de nuestros colegas de Caracas, pero enmarcado en una perspectiva nacional, y el hecho de tener la ciudad de Mérida una intensa y variada vida cultural y científica, nos motivó a cambiar el nombre del ciclo de conferencias por el **COLOQUIO MATEMATICO DE MERIDA**, manifestando así nuestra intención e interés de insertarlo en la programación científica y cultural de la ciudad. Así procuramos desarrollar un espacio de confrontación de ideas, alrededor de la experiencia matemática, optimizando los recursos que la Nación invierte para el avance de la ciencia y la cultura.

Con la finalidad de testimoniar esta labor, el Profesor Oswaldo Araujo propuso la idea de publicar en Notas de Matemática, aquellas conferencias previamente autorizadas por sus autores. Este ejemplar, materialización de esa idea, consta de cinco conferencias dictadas entre el 3 de Diciembre de 1992 y el 17 de Enero de 1994. Pensamos que su publicación puede ser un estímulo para los futuros conferencistas.

Aprovechamos la ocasión para agradecerle a los Profesores Jorge Vargas, Martin Jones, Wilfredo Urbina, Carlos Benítez y Richard Gundy, su colaboración al autorizarnos la publicación de sus conferencias, y a la Comisión de Publicaciones del Departamento de Matemáticas de la Facultad de Ciencias de la ULA por apoyar e instrumentar esta iniciativa.

Finalmente, expresamos nuestra convicción de que esta labor de documentar una actividad matemática solamente podrá tener continuidad si este esfuerzo encuentra eco en nuestra comunidad.

Los editores

Oswaldo Araujo
Secretario General de la AMV
Capítulo Mérida

Victor Padrón
Coordinador del Posgrado
de Matemáticas

BOSQUEJO DE LA EVOLUCION DE LA TEORIA DE LIE

JORGE VARGAS

Facultad de Matemática, Astronomía y Física

Universidad Nacional de Córdoba.

Mérida, 3 de Diciembre de 1992.-

Bosquejo de la evolución de la teoría de Lie

Jorge Vargas
Facultad de Matemática, Astronomía y Física [1]
Universidad Nacional de Córdoba
Argentina

En algún lugar de la historia aprendimos a resolver equaciones del tipo $x^2 + bx + c = 0$, logrando para las soluciones la expresión explícita

$$(1/2)(-b \pm \sqrt{b^2 - 4c}) \tag{1}$$

Con la introducción, por Vieta en 14.. de la notación exponencial, Cardano, Rufini y otros descubrieron que las soluciones de una equación polinómica de tercer o cuarto grado se obtienen operando con suma, producto y radicación sobre los coeficientes de la ecuación y raíces de la unidad.

El problema de calcular las raíces de equaciones polinómicas de grado mayor o igual a cinco, tiene un punto crucial al comienzo del siglo pasado con los trabajos de Galois. Galois, en lugar de buscar fórmulas explícitas del tipo (1) nos provee de un nuevo y explícito modo, que permite, para una ecuación polinómica

$$\sum_{i=1}^{n} a_i x^i = 0 \tag{2}$$

decidir si sus raíces pueden ser calculadas a partir de los coeficientes de (2) y raíces de la unidad por medio de las operaciones de suma, producto y radicación.

Para describir la técnica de Galois, por simplicidad, supondremos que los coeficientes a_i de (2) son números racionales. Recordemos que Gauss, unos años antes, había probado que existen n números complejos $z_1, \cdots, z_n$ que resuelven la ecuación (2). Por cierto, este teorema de Gauss es netamente existencial, no provee ningún método de cálculo, probablemente, este fue uno de los primeros teoremas que prueban existencia de un objeto matemático sin explícitamente construirlo.

[1] Conferencia dictada en Ciencias-U.L.A., Mérida, Venezuela, Nov. 1992

Sea F igual al menor subcuerpo del cuerpo de los números complejos C que contiene a $z_1, \cdots, z_n$.

Por ejemplo, si la ecuación es $x^2 - 3 = 0$, entonces

$$F = \{a + b\sqrt{3}, \quad \text{con } a, b, \text{racionales}\}$$

Sea

$$G_f = \{\sigma \,:\, F \to F \text{ es morfismo de cuerpos}\}$$

$$= \{\sigma \,:\, F \to F \text{ tal que}$$

$$\sigma(x+y) = \sigma(x) + \sigma(y),\ \sigma(xy) = \sigma(x)\sigma(y)\,\forall x, y \in F\}$$

Se tiene,

Teorema (Galois). Es posible calcular $z_1, \cdots, z_n$ operando con suma, producto y radicación en $a_0, \cdots, a_{n-1}$ y números racionales sí, y sólo sí el grupo G_f es soluble.

Recordemos que un grupo H se dice soluble, si su serie derivada se estaciona en la identidad.

Por otro lado, desde el advenimiento de las ecuaciones diferenciales surgió el problema del cálculo explícito de sus soluciones, esto es, encontrar expresiones explícitas para el conjunto de soluciones de una ecuación diferencial dada, en lenguaje más clásico, el problema se expresa en: encontrar "la integral general" de una ecuación diferencial. En particular, para ecuaciones diferenciales ordinarias lineales de primer orden

$$x' = a(t)x + b(t) \tag{3}$$

hemos aprendido a calcular la expresión explícita de sus soluciones operando algebraicamente, integrando o exponenciando funciones. En efecto, la solución general de (3) es

$$x(t) = \int b e^{-\int a}$$

Por otro lado, Lagrange, para ecuaciones diferenciales lineales ordinarias probó que si se conoce explícitamente una solución, entonces es posible transformar la ecuación en una de un orden menor, en otras palabras descubrió el método de reducción de orden, por lo tanto, cada vez que conocemos una solución de una ecuación diferencial lineal, ordinaria podemos reducir el orden y así, con suerte, llegar a una ecuación de primer orden. Esta técnica, tiene un pequeño inconveniente: ¡sólo es posible calcular las soluciones si conocemos las soluciones!
Lie en la segunda mitad del siglo pasado descubrió un análogo al teorema de Galois y por consiguiente, en principio, es posible decir si las soluciones de un sistema de equaciónes diferenciales lineales, ordinario de primer orden [2]

$$x' = A(t)(x(t)) + b(t) \tag{4}$$

$$\text{donde } A : R \rightarrow R^{n \times n}, \quad b : R \rightarrow R^n \text{ son diferenciables}$$

son calculables por medio de las operaciones de suma, producto, cálculo de primitivas y/o exponenciación de funciones a partir de los coeficientes de $A(t), b(t)$.

La idea de Lie consiste, como lo hizo Galois, en asociar a la ecuación (4) un grupo, hoy en día llamado grupo de Lie de la ecuación. Para definir este grupo, pensemos las soluciones de (4) como curvas en R^{n+1}. De modo intuitivo, el grupo de Lie de la ecuación (4) es el conjunto de difeomorfismos de R^{n+1} que transforma cada curva solución de (4) en otra curva solución de (4).

Por ejemplo, para la ecuación $x' = x$ (aquí $n = 1$), las curvas solución son $x_c(t) = (t, ce^t), \quad c \in R$. Un difeomorfismo de R^2 es $f_s(a, b) = (a + s, b)$, calculemos $f_s(t, ce^t) = (t + s, ce^t e^s) = (t + s, ce^t) = (u, ce^{u-s}) = (u, ce^{-s}e^u)$, en la penúltima igualdad hicimos $t + s = u$, por lo tanto, $f_s(t, ce^t)$ es otra solución de la ecuación $x' = x$ y en consecuencia, f_s pertenece al grupo de Lie de la ecuación $x' = x$.

Formalmente, el grupo de Lie lo definimos, a partir de

$$\Sigma = \{(t, x(t))_{t \in R}, \text{ donde } x : \ R \rightarrow R^n, \text{ es solucion de (4) } \}$$

[2] R = números reales

Considerando las familias $(f_s)_{s\in R}$ de difeomorfismos de R^{n+1} que satisfacen:

$$i)\ f_{s+r} = f_s \circ f_r \quad s,r, \in R$$

$$\text{ii)}\ f_0 = \text{función identidad}$$

$$iii)\ \text{Para cada curva } (t,x(t))_{t\in R} \in \Sigma, \text{ la curva, } f_s(t,x(t)) \text{ es del tipo}$$

$$(u,y(u))_{u\in R} \text{ y la función } y(u) \text{ es solución de (4)}$$

$$\text{iv) la función } (s,a) \to f_s(a) \text{ es diferenciable}$$

Por definición el *grupo de Lie* de la ecuación (4) es el grupo de difeomorfismos de R^{n+1} generado por todos los elementos f_s de todas las familias $(f_s)_{s\in R}$ que satisfacen i), ii), iii) y iv).
Convengamos en denotar dicho grupo por G_e
Observemos que si conocemos una solución $x(t)$ de (4) y un elemento $f \in G_e$, entonces tenemos otra solución de (4), a saber $f(x(t))$.

Un teorema de Lie, expresa: Es posible calcular todas las soluciones de (4) por medio de las operaciones de suma, producto, exponenciación y cálculo de primitivas realizadas en los coeficientes de $A(t), b(t)$ si G_e es un grupo soluble cuya dimensión es mayor o igual a n.

Para completar el enunciado del teorema de Lie debemos decir que significa dimensión de G_e y como calcular G_e.
Otra gran contribución de Lie, consistió en que algebrizó tanto el problema de calcular $dimG_e$ como el de conocer cuando G_e es un grupo soluble.
Para esto procedió de la manera siguiente:
Para cada familia f_s que satisface i), ii), iii) y iv) asocia una función

$$X_f \ : \ R^{n+1} \to R^{n+1}$$

por la fórmula

$$X_f(x_1,\ldots,x_{n+1}) = \left(\frac{dh_1(s,x_1,\ldots,x_{n+1})}{ds}_{|s=0},\ldots,\frac{dh_{n+1}(s,x_1,\ldots,x_{n+1})}{ds}_{|s=0}\right)$$

si $f_s(x_1,\ldots,x_{n+1}) = (h_1(s,x_1,\ldots,x_{n+1}),\ldots,h_{n+1}(s,x_1,\ldots,x_{n+1}))$

con h_j funciones de R^{n+2} en R.

Por último, consideró

$$\mathcal{L}_e := \{X_f \text{ tal que } X_f \text{ proviene de una familia } f_s\ \}$$

entonces:
a) Lie probó que $\mathcal{L}_e$ es un espacio vectorial de dimensión finita; a $dim\mathcal{L}_e$ lo denominó la dimensión de G_e.
b) Para cada $X = (\psi_i)_{1\leq i\leq n+1}$, $Y = (\eta_i)_{1\leq i\leq n+1} \in \mathcal{L}_e$, Lie definió una nueva función $[X, Y]$ de R^{n+1} en sí mismo por la fórmula,

$$[X,Y] = (\sum_j (\psi_j \frac{\partial}{x_j}(\eta_i) - \eta_j \frac{\partial}{x_j}(\psi_i))_{1\leq i\leq n+1}$$

y demostró que $[X, Y] \in \mathcal{L}_e$.
c) Escribamos cada función $X(t, y_1, \ldots, y_n)$ de R^{n+1} en sí mismo como

$X = (\psi, \eta)$ donde $\psi\ :\ R^{n+1} \to R,\quad \eta\ :\ R^{n+1} \to R^n,$

Sea $g(t, y_1, \ldots, y_n) := A(t)(y_1, \ldots, y_n) + b(t) = (g_1, \ldots, g_n)$.

Lie verificó

$X \in \mathcal{L}_e$ sí, y sólo sí vale la igualdad

$$\partial_t(\eta) + \partial_y(\eta)(g(t,y)) - \partial_t(\psi).g(t,y) - (g(t,y).\partial_y(\psi))\, g(t,y) =$$

$$= \psi\partial_t(g) + (\partial_y g)(\eta(t,y)).$$

d) Tambien demostró, G_e es un grupo soluble sí, y sólo sí el álgebra $(\mathcal{L}_e, [\,,\,])$ es soluble.
Recordemos que para un álgebra $(g, [\,,\,])$ se define su serie derivada inductivamente por $\mathcal{D}^1(g)$ = subespacio generado por $\left\{_{[x,y],x,y\in g\}}\right\}$, y $\mathcal{D}^k(g) := \mathcal{D}^1(\mathcal{D}^{k-1}(g))$, por definición, el álgebra $(g, [\,,\,])$ es soluble si $\mathcal{D}^k(g) = 0$ para algún k.
e) Un problema abierto en esta dirección es producir fórmulas explítas para la $dim\mathcal{L}_e$ en función de g, por cierto, hay cotas superiores, consultar (Ov, pág. 99)
f) Por el teorema de Peanno-Picard (PP) se tiene que el conocimiento de $\mathcal{L}_e$

implica el conocimiento teórico de G_e, puesto que que para cada $X \in \mathcal{L}_e$, $p \in R^{n+1}$ PP nos dice que existe una única curva $\gamma_{X,p}$ en R^{n+1} tal que $\dot{\gamma_{X,p}} = X(\gamma_{X,p})$, $\gamma_{X,p}(0) = p$ y es fácil deducir que $f_s(p) := \gamma_{X,p}(s)$ satisface i), ii), iii) y iv)

En un arduo trabajo, Lie desarrollo una teoría similar para ecuaciones ordinarias o en derivadas parciales tanto lineales como no lineales; Blumen y Kumei, en los ochenta retomaron estas investigaciones probando que un sistema de ecuaciónes en derivadas parciales es equivalente a un sistema lineal si el grupo de la ecuación satisface ciertos requerimientos. Además, probaron que si el grupo de Lie de un sistema de ecuaciónes lineales en derivadas parciales es de cierto tipo, entonces el sistema es equivalente a uno lineal a coeficientes constantes, consultar (B-K).

A partir de los trabajos de Dirac y otros desde la década del treinta se investigó en problemas del tipo siguiente:
Sea G un grupo de Lie actuando en una variedad differenciable M, de esto, tenemos un morfismo de grupos $\bullet M \to Diff(M)$ donde se verifica que la función $G \times M \to M$, $(g, m) \to g \bullet (m)$ es diferenciable.
Para cada función $f : M \to R$, definimos una nueva función $g \bullet f$ por la igualdad

$$(g \bullet f)(m) = f(g \bullet (m)) \; (m \in M, \, g \in G)$$

Sea $D : C^\infty(M) \to C^\infty(M)$ un operador diferencial que conmuta con la acción de G, esto es,

$$D(g \bullet f) = g \bullet (D(f)) \; (g \in G, \, f \in C^\infty(M))$$

Para cada número complejo λ sea $C^\infty(M)_\lambda =$ el autoespacio de autovalor λ de D, por tanto,

$$C^\infty(M)_\lambda = \{f \in C^\infty(M) \, \text{tal que} \, D(f) = \lambda f\}$$

Un ejercicio sencillo permite verificar que si $f \in C^\infty(M)_\lambda$ entonces $D(f) \in C^\infty(M)_\lambda$.

Un ejemplo concreto de esta situación es $M = R^n$, $G = Iso(R^n)$ =grupo de isometrías de R^n, y $D =$ laplaciano, se tiene:

Teorema (Helgason, 1970). Sí $\lambda \neq 0$, para cada $f \in C^\infty(R^n)_\lambda$ no nula fija, el subespacio generado por

$$\{g \bullet f,\ g \in Iso(R^n)\ \}$$

es denso en $C^\infty(R^n)_\lambda$ munido de la topología de convergencia uniforme en compactos para las funciones y todas sus derivadas parciales.

Notar que el resultado no es cierto para $\lambda = 0$, puesto que la función constante igual a 1 pertenece a $C^\infty(R^n)_0$ y $g \bullet 1 = 1\ \forall\, g \in Iso(R^n)$.

En la bola unidad B_n de R^n, Atiyah, Hirzebruck y otros construyeron una familia de operadores diferenciales lineales, elípticos parametrizados por un conjunto Γ, digamos $(D_\mu)_{\mu\in\Gamma}$ (los operadores de Dirac generalizados), estos operadores conmutan con la acciones generalizadas $\bullet_\mu$ del grupo de isometrías noeuclideanas de B_n. Atiyah-Schmid probaron que para el núcleo en L^2 de cada D_μ vale un resultado análogo al de Helgason. En un trabajo a aparecer en el Trans. AMS, conjuntamente con E. Galina hemos descripto los autoespacios en L^2 de un D_μ fijo en términos de los núcleos en L^2 de otros D_σ.

El autor desea agradecer al Dr. O. Araujo por comentarios sobre la presente nota y al grupo de álgebra de la ULA por la invitación a participar de sus actividades.

Bibliografía

Atiyah-Schmid, A geometric construction of the discrete series for semisimple Lie groups, Inv. Math. 1977, Vol. 42

Blumen-Kumei, Symmetries and differential equations, Springer Verlag, 1991

Galina-Vargas, Eigenvalues and eigenvectors for the twisted Dirac operator pver $SU(n,1)$ and $SPIN(2n,1)$, a aparecer en Trans. AMS , ICTP preprint 91/347

Ovsiannikov, Group analysis of differential equations, Academic Press, 1982

Dirección: FAMAF, Ciudad Universitaria, 5016 Córdoba, Argentina

PROBLEMAS DE LOS PROFETAS EN LA TEORIA DE LOS TIEMPOS OPTIMOS

MARTIN L. JONES

Departamento de Matemáticas

Universidad de Charleston

Mérida, 17 de Mayo de 1993.-

Problemas de los Profetas en la Teoría de los Tiempos Óptimos

Martin L. Jones
Departamento de Matemáticas
Universidad de Charleston, South Carolina
Charleston, South Carolina 29424
EEUU

Introducción Supóngase que una persona observa los eventos de una serie de juegos de azar. En cualquier momento, esta persona puede dejar de observar y recibir como premio la cantidad observada en el último juego. Nunca tiene la oportunidad para regresar y recoger los premios observados en los juegos en "el pasado". Sin embargo, sabe la distribución de los juegos en "el futuro" antes de observarlos. Este jugador quiere encontrar una "regla de parar", la cual puede ayudarlo con su decisión sobre el tiempo más apropriado para terminar las observaciones y recoger su premio. El objetivo clásico es escoger una regla que maximice la esperanza del premio recibido.

Además, supóngase que hay otra persona con la libertad de recoger premios de los juegos ya observados. Esta persona, por supuesto, recibirá un premio esperado más grande que el jugador que usa las reglas de parar. El propósito de este artículo es introducir al lector a unos de los problemas sobre comparaciones entre los premios esperados de estos dos tipos de observadores.

Con más exactitud, sean $X_1, X_2, X_3, \ldots$ una serie de variables aleatorias definidas en un espacio de probabilidades (Ω, F, P), y sean $F_1, F_2, F_3, \ldots$ una serie de σ-fields aumentativos tales que para cada $k = 1, 2, 3, \ldots$, X_k es F_k mensurable. Generalmente, F_k es la σ-algebra de Borel generado por las primeras k variables aleatorias en la serie. Una variable aleatoria, τ, definida en (Ω, F, P) con valores en el conjunto $\{1, 2, 3, \ldots, +\infty\}$ se llama una "regla de parar" si para cada $k = 1, 2, 3, \ldots$, $\{\tau = k\} \in F_k$.. La variable X_τ es igual a X_k en el conjunto $\{\tau = k\}$ y cero en $\{\tau = +\infty\}$. A menudo, esta variable se llama "la serie terminada por τ". La condición $\{\tau = k\} \in F_k$ indica que la decisión de terminar las observaciones puede basarse solamente en las observaciones en el presente y en el pasado, pero no en las del futuro. El "valor" de la serie $X_1, X_2, X_3, \ldots$ se indica por $V(X_1, X_2, X_3, \ldots)$ y se define por $V(X_1, X_2, X_3, \ldots) = \sup_\tau EX_\tau$, donde el supremum es sobre todas las reglas, τ, tales que $P\{\tau < +\infty\} = 1$ y

EX_τ existe. Una pregunta obvia es la siguiente: ¿Cuándo existe una regla óptima y cuáles son las condiciones necesarias? En otras palabras, ¿cuándo existe una regla, τ^*, tal que $V(X_1, X_2, X_3, \ldots) = \sup_\tau EX_\tau = EX_{\tau^*}$? La respuesta está fuera del alcance de este artículo (ver [1]), pero en el problema de "horizontes finitos", cuando la serie de juegos es finita, una regla óptima existe y se puede calcular, por lo menos en teoría. Se refiere al proceso que se usa para determinar la regla óptima y el valor de la serie finita se refiere como "el principio de programación dinámica" o "indución hacia atrás" (otra vez, ver [1]). Se basa en una definición de recursión, en una manera hacia atrás, con las variables aleatorias

$$\gamma_n = X_n,$$
$$\gamma_{n-1} = \max\{X_{n-1}, E(\gamma_n \mid F_{n-1})\},$$
$$\vdots$$
$$\gamma_1 = \max\{X_1, E(\gamma_2 \mid F_1)\}.$$

Después se puede demostrar que la regla óptima, τ, está dada por $\tau = \min\{k \geq 1 : X_k = \gamma_k\}$.

Ejemplo 1. Supóngase que los eventos de una serie de cuatro variables aleatorias X_1, X_2, X_3, y X_4 son los números observados cuando se dan vueltas a las flechas de las ruletas independientes en el dibujo.

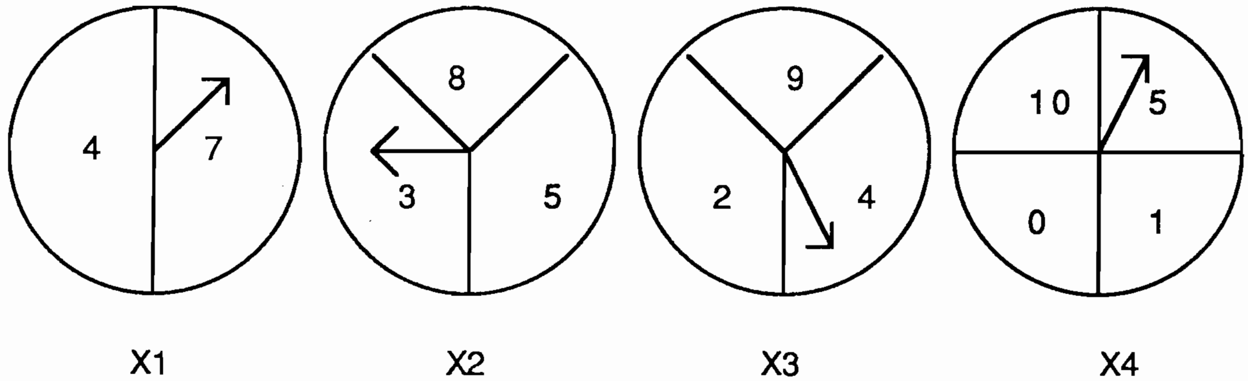
Se puede usar el principio de programación dinámica y la independencia de las variables aleatorias para calcular el valor de esta serie, $V(X_1, X_2, X_3, X_4) = \frac{121}{18} = 6.72222$ y demostrar que la regla óptima, τ, es

$\tau = 1$, cuando $X_1 = 7$,

$\tau = 2$, cuando $X_1 \neq 7$, y $X_2 = 8$,

$\tau = 3$, cuando $X_1 \neq 7$, $X_2 \neq 8$, y $X_3 = 9$,

$\tau = 4$, cuando $X_1 \neq 7$, $X_2 \neq 8$, y $X_3 \neq 9$.



Problemas de los Profetas Regresando a las ideas presentadas en los primeros párrafos de la introducción, considere ahora el premio esperado por la segunda persona que está observando la serie, el profeta. A causa del hecho de que este observador puede escoger premios ya observados, su premio esperado está dado por $M(X_1, X_2, X_3, \ldots)$ donde $M(X_1, X_2, X_3, \ldots) = E(\sup\{X_1, X_2, X_3, \ldots\})$. En el caso de horizontes finitos, $M(X_1, X_2, \ldots, X_n) = E(\max\{X_1, X_2, \ldots, X_n\})$. En el Ejemplo 1, se puede demostrar que $M(X_1, X_2, \ldots, X_n) = 8.02778$, y la diferencia entre las esperanzas de los dos observadores está dada por $M(X_1, X_2, \ldots, X_n) - V(X_1, X_2, \ldots, X_n) = 1.30556$. El objeto es determinar la constante universal, d_n, tal que $M(X_1, X_2, \ldots, X_n) - V(X_1, X_2, \ldots, X_n) \leq d_n$ sobre todas las distribuciones para $X_1, X_2, \ldots, X_n$ en una clase predeterminada, C, de distribuciones. Esta desigualdad, cuando se establezca, se llama una "desigualdad de profetas". También es posible comparar las razones de las cantidades $M(X_1, X_2, \ldots, X_n)$ y $V(X_1, X_2, \ldots, X_n)$ y construir desigualdades de un tipo diferente. Unos ejemplos de desigualdades de profetas aparecen en los Teoremas 1, 2, y 3 abajo.

Teorema 1 (Hill y Kertz [6]) Sean $X_1, X_2, X_3, \ldots$ una serie de variables aleatorias independientes y con valores en el intervalo [a, b] donde a < b son constantes finitos. Entonces $M(X_1, X_2, \ldots, X_n) - V(X_1, X_2, \ldots, X_n) \leq \frac{b-a}{4}$.

Es interesante que la constante, d_n, en este caso no depende de n, y la diferencia máxima se puede realizar con una serie de solamente dos variables aleatorias, como en el ejemplo siguiente. En particular, cuando el intervalo es $[0, 1]$ la constante, d_n, es $\frac{1}{4}$.

Ejemplo 2 Sean X_1 y X_2 variables aleatorias independientes donde $X_1 = \frac{a+b}{2}$ y X_2 toma los valores $\{a, b\}$ cada uno con probabilidad $\frac{1}{2}$. Cálculos fáciles indican que $M(X_1, X_2, \ldots, X_n) - V(X_1, X_2, \ldots, X_n) = \frac{b-a}{4}$.

Teorema 2 (Hill y Kertz [7]) Sean $X_1, X_2, X_3, \ldots$ una serie de variables aleatorias arbitrarias que toman valores en el intervalo $[0, 1]$. Entonces, para cada $n \geq 2$,

$$M(X_1, X_2, \ldots, X_n) - V(X_1, X_2, \ldots, X_n) \leq \left(\frac{n-1}{n}\right)^n.$$

Observe que en el Teorema 2, las constantes, d_n, dependen de n y son más grandes que $\frac{1}{4}$ cuando $n > 2$. La distribución extrema en este problema es una martingala (ver [7]).

Teorema 3 (Jones [8]) Sean $X_1, X_2, X_3, \ldots$ una serie de variables aleatorias independientes que toman valores en el intervalo $[0, 1]$ y sea $c > 0$ una constante. Para cada $k = 1, 2, 3, \ldots$ defina $Y_k = X_k - kc$. Entonces, para cada $n \geq 2$,

$$M(Y_1, Y_2, \ldots, Y_n) - V(Y_1, Y_2, \ldots, Y_n) \leq \left(\frac{n-1}{n}\right)^n.$$

En el Teorema 3, cada observador necesita pagar la cantidad c por cada observación que quiere ver. Es interesante notar que, aunque las variables aleatorias son independientes en este caso, las constantes, d_n, son exactamente las mismas como las del Teorema 2 donde las variables aleatorias son arbitrarias, pero no hay un costo de observación.

Aunque hay muchos más ejemplos conocidos de desigualdades de profetas, hay muchos problemas de este tipo que quedan sin resolver. Dos problemas no resueltos aparecen abajo. El primero es una mezcla de los teoremas 2 y 3. Mientras las mejores constantes en los casos de variables arbitrarias (Teorema 2), y variables independientes con un costo de observación (Teorema 3) son conocidas, el problema siguiente no está resuelto.

Problema 1 Sean $X_1, X_2, X_3, \ldots$ una serie de variables aleatorias arbitrarias que toman valores en el intervalo [0, 1] y sea $c > 0$ una constante. Para cada $k = 1, 2, 3, \ldots$ defina $Y_k = X_k - kc$. Para cada $n \geq 2$, ¿cuál es la mejor constante, d_n, tal que

$$M(Y_1, Y_2, \ldots, Y_n) - V(Y_1, Y_2, \ldots, Y_n) \leq d_n?$$

El otro problema, de un tipo diferente, que no está resuelto todavía, trata sobre una variación en los premios recibidos por los observadores. Aquí los observadores reciben el promedio de las observaciones en el momento en el cual deciden terminar sus observaciones.

Problema 2 Sean $X_1, X_2, X_3, \ldots$ una serie de variables independientes que toman valores en el intervalo [0, 1]. Para cada $k = 1, 2, 3, \ldots$ defina $Y_k = \dfrac{X_1 + X_2 + \ldots + X_k}{k}$. Para cada $n \geq 2$, ¿cuál es la mejor constante, d_n, tal que

$$M(Y_1, Y_2, \ldots, Y_n) - V(Y_1, Y_2, \ldots, Y_n) \leq d_n?$$

Problemas de los Profetas Permutando el Orden de las Observaciones Supóngase otra vez que hay dos observadores que están observando una serie finita de variables aleatorias $X_1, X_2, \ldots X_n$. Uno es ordinario y debe usar reglas de parar no anticipadas para decidir el momento en que debe dejar de observar, mientras otro es un profeta que puede "ver el futuro" y escoger el máximo de la serie de observaciones. En esta situación, sin embargo, se permite al primer jugador la oportunidad de escoger el orden de la serie antes de empezar las observaciones. De esta manera puede mejorar los premios esperados después de considerar las distribuciones de las variables. Con más exactitud, sea $\pi \in P_n$, el conjunto de todas las permutaciones del conjunto $\{1, 2, \ldots, n\}$ y defina $W(X_1, X_2, \ldots, X_n)$, por

$$W(X_1, X_2, \ldots, X_n) = \max\{V(X_{\pi(1)}, X_{\pi(2)}, \ldots, X_{\pi(n)}) : \pi \in P_n\}.$$

Por ejemplo, se puede demostrar en el Ejemplo 1 que el mejor orden es X_4, X_3, X_2, X_1, y que $W(X_1, X_2, \ldots, X_n) = 7.91667$. Como antes, se puede hacer comparaciones entre los premios esperados de jugadores de tipos diferentes: profetas, jugadores ordinarios con el privilegio de cambiar el orden de la serie, y jugadores ordinarios sin este privilegio. Mientras hay varios artículos enfocando en este tipo de problema, hay muchas cosas desconocidas (ver [2], [3], [4] y [5]). El teorema siguiente de Gilat [2], trata el caso de dos variables aleatorias independientes que toman valores en el intervalo [0, 1].

Teorema 4 (Gilat [2]) Sean X_1 y X_2 dos variables aleatorias independientes que toman valores en el intervalo [0, 1]. Entonces

$$M(X_1, X_2) - W(X_1, X_2) \leq \frac{5\sqrt{5} - 11}{2}.$$

Es interesante considerar la distribución extrema en este problema, construida de la manera siguiente. Sea $w^* = \frac{\sqrt{5} - 1}{2}$, llamada la "Razón Dorada", y sean X_1 y X_2 variables aleatorias independientes tales que $P(X_1 = 1 - w^*) = w^* = 1 - P(X_1 = 1)$, y $P(X_2 = w^*) = w^* = 1 - P(X_2 = 0)$. Observe que las variables aleatorias son de "orden indiferente", es decir $W(X_1, X_2) = V(X_1, X_2) = V(X_2, X_1)$, aunque no tienen distribuciones idénticas.

Las mejores constantes en los casos $n > 2$ no son conocidas, sin embargo Gilat hizo conjecturas que indican que las distribuciones extremas consistirán de variables que son de orden indiferente, pero que no tienen distribuciones idénticas. También Gilat cree que las constantes aumentarán con n.

Referencias

1. Chow, Y.S., et. al., (1971) Great Expectations: The Theory of Optimal Stopping, Houghlin Mifflin, Boston, Mass.

2. Gilat, D., (1985) "A Prophet Inequality with Order Selection for Two Independent Random Variables", Unpublished Manuscript, Tel Aviv University, Israel.

3. Gilat, D., (1987) "On the Best Order of Observation in Optimal Stopping Problems", Journal of Applied Probability, **24**, 773-778.

4. Hill, T.P., (1983) "Prophet Inequalities and Order Selection in Optimal Stopping Problems", Proceedings of the American Mathematical Society, **88**, No. 1, 131-137.

5. Hill, T.P., and Hordijk, A., (1985) "Selection of Order of Observation in Optimal Stopping Problems", Journal of Applied Probability, **22**, 177-184.

6. Hill, T.P., and Kertz, R.P., (1981) "Additive Comparisons of Stop Rule and Supremum Expectations of Uniformly Bounded Independent Random Variables", Proceedings of the American Mathematical Society, **83**, 582-585.

7. Hill, T.P., and Kertz, R.P., (1983) "Stop Rule Inequalities for Uniformly Bounded Sequences of Random Variables", Transactions of the American Mathematical Society, **278**, 197 - 207.

8. Jones, M. (1990) "Prophet Inequalities for Optimal Stopping Problems with a Cost of Observation", Journal of Multivariate Analysis, **34**, 238 - 253.

(DES) IGUALDADES DEL PARALELOGRAMO, DEL ROMBO, DEL RECTANGULO Y DEL CUADRADO. GENERALIZACION DE CARLSSON

CARLOS BENITEZ

Departamento de Matemáticas

Universidad de Extremadura

Mérida, 23 de Septiembre de 1993.-

(DES)IGUALDADES DEL PARALELOGRAMO, DEL ROMBO, DEL RECTANGULO Y DEL CUADRADO. GENERALIZACION DE CARLSSON

Carlos Benítez

Dpto. de Matemáticas. Univ. de Extremadura.

06071. Badajoz. España.

Sea E un espacio normado sobre el cuerpo $\mathbb{K}$ ($\mathbb{R}$ o $\mathbb{C}$), con $S=\{x\in E: \|x\|=1\}$ como esfera unidad.

En 1935 Jordan y von Neumann [11] probaron que la norma de E está inducida por un producto escalar (E es prehilbertiano, o pH) si y solo si se verifica la **igualdad del paralelogramo**,

$$\boxed{\forall x,y\in E, \quad \|x+y\|^2+\|x-y\|^2=2(\|x\|^2+\|y\|^2)} \tag{1}$$

La demostración de la suficiencia de tal condición (la necesidad es evidente) es puramente analítica y consiste en ver que,

$$(x|y)=\tfrac{1}{4}(\|x+y\|^2-\|x-y\|^2), \quad \text{si } \mathbb{K}=\mathbb{R},$$

$$(x|y)=\tfrac{1}{4}[(\|x+y\|^2-\|x-y\|^2)+ i(\|x+iy\|^2-\|x-iy\|^2)], \quad \text{si } \mathbb{K}=\mathbb{C},$$

son productos escalares en E que inducen su norma.

Veámoslo para $\mathbb{K}=\mathbb{R}$. Es inmediato que $\|x\|^2=(x|x)$ y que $(x|y)=(y|x)$, y se sigue de la hipótesis que $(x+y|z)=(x|z)+(y|z)$. Por consiguiente, $(2x|y)=2(x|y)$, y reiterando el razonamiento, $(\lambda x|y)=\lambda(x|y)$, para todo $\lambda\in\mathbb{N}$. Por otra parte como $(-x|y)=-(x|y)$, también aquello es cierto para $\lambda\in\mathbb{Z}$, y, en definitiva (poniendo $\lambda^{-1}x$ donde pone x), para todo $\lambda\in\mathbb{Q}$. Finalmente, como la función $\lambda\in\mathbb{R}\longrightarrow(\lambda x|y)-\lambda(x|y)$ es continua y se anula en $\mathbb{Q}$, se llega a que $(\lambda x|y)=\lambda(x|y)$, para todo $\lambda\in\mathbb{R}$.

Como corolario se obtiene que E es pH si y solo si lo son sus subespacios bidimensionales, y también que un espacio normado complejo es pH si y solo si lo es cuando se considera como real. Tal hecho y la naturaleza de las cuestiones que vamos a tratar, permiten suponer desde ahora que <u>E es el plano real $\mathbb{R}^2$</u>.

En 1947 Day [7] probó que era suficiente que aquella igualdad se verificara para vectores de la misma norma. Es decir que E es pH si y solo si se verifica la **igualdad del rombo**,

$$\boxed{\forall x,y\in S, \quad \|x+y\|^2+\|x-y\|^2=4} \tag{2}$$

A diferencia de la anterior la elegante demostración de Day es de naturaleza geométrica y se basa en el hecho de que, cualquiera que sea la curva convexa y simétrica respecto del origen, S, existen elipses inscritas y circunscritas a S también simétricas respecto al origen, que la tocan en, al menos, cuatro puntos (dos linealmente independientes y sus respectivos opuestos). Concretamente, utiliza un resultado de Loewner que dice que entre todas las elipses inscritas (circunscritas) a S, existe una y solo una cuya área es máxima (mínima) y a la que ocurre aquello.

Eso supuesto, sea C, por ejemplo, una elipse inscrita a S que la toca en dos puntos linealmente independientes y sea $|\ |$ la norma de E cuya esfera unidad es C. Si $S=C$ no hay nada que probar (S es una elipse y E ya es pH), y si $S\neq C$ basta tomar dos puntos $x,y\in S\cap C$ tales que en el menor de los arcos de S que los unen no haya otro punto de $S\cap C$, para obtener,

$$\|x\|=|x|=\|y\|=|y|=1\ ,\quad \|x+y\|<|x+y|\ ,\quad \|x-y\|\leqslant|x-y|\ ,$$

y, por tanto, la contradicción,

$$4=\|x+y\|^2+\|x-y\|^2<|x+y|^2+|x-y|^2=4\ .$$

Posteriormente, Schoenberg en 1952 [13] llamó la atención sobre el hecho de que la demostración de Day lo que decía era que E es pH si y solo si se verifica la **desigualdad del rombo**,

$$\boxed{\forall x,y\in S,\quad \|x+y\|^2+\|x-y\|^2\sim 4} \tag{3}$$

donde $\sim$ denota alguno de los signos $\leqslant$ ó $\geqslant$.

En esa misma línea, Benítez y del Río probaron en 1984 [3] que E es pH si y solo si se verifica la **desigualdad del rectángulo**,

$$\boxed{x,y\in E,\quad x\perp y\quad \Longrightarrow\quad \|x+y\|^2+\|x-y\|^2\sim 2(\|x\|^2+\|y\|^2)} \tag{4}$$

donde $x\perp y$ significa que x es ortogonal a y en el sentido de Birkhoff [4, 10], es decir, $\|x\|\leqslant\|x+\lambda y\|$, para todo $\lambda\in\mathbb{R}$.

La demostración se basa en una modificación de un resultado de Day [8], que dice que E es pH si y solo si,

$$\forall x,y\in S,\ \exists\,\alpha,\beta\in\mathbb{R}_+\ :\ \|\alpha x+\beta y\|^2+\|\alpha x-\beta y\|^2\sim 2(\alpha^2+\beta^2),$$

y en otro de Benítez [2], según el cual,

$$\forall x,y\in E,\ \exists\,\delta\in\mathbb{R}_+\ :\ x+\delta y\perp x-\delta y.$$

Al igual que hizo Schoenberg con la igualdad del rombo, Amir en 1986 ([1], pág. 50) llamó la atención sobre el hecho de que aquella demostración lo que dice es que E es pH si y solo si se verifica,

$$\boxed{x, y \in E, \quad x \approx y \quad \Longrightarrow \quad \|x+y\|^2 + \|x-y\|^2 \sim 2(\|x\|^2 + \|y\|^2)} \tag{5}$$

donde $\approx$ es una relación binaria en E que "admite diagonales", es decir, tal que

$$\forall x, y \in E, \ \exists \delta \in \mathbb{R}_+ \ : \ x + \delta y \approx x - \delta y.$$

(Merece la pena señalar que todas las ortogonalidades en espacios normados conocidas por el redactor de estas líneas, excepto la muy restrictiva de Roberts, son relaciones binarias que admiten diagonales. Es decir puede suponerse en (4) que $\perp$ es una cualquiera de tales ortogonalidades).

Una pregunta que surge de inmediato es la de si también caracterizará a los espacios prehilbertianos la **desigualdad del cuadrado**,

$$\boxed{x, y \in S, \quad x \approx y \quad \Longrightarrow \quad \|x+y\|^2 + \|x-y\|^2 \sim 4} \tag{?}$$

Aunque esta cuestión permanece abierta, pueden darse algunas respuestas parciales relacionadas con el caso posiblemente más interesante, que es aquel en que $\approx$ representa a la ortogonalidad de Birkhoff, $\perp$, (ver [3]):

(i) S es una elipse si y solo si $\|x+y\|^2 \sim 2$, para todo par $x, y \in S$ tal que $x \perp y$.

(ii) Cualquiera que sea S, existen $x, y \in S$ tales que $x \perp y$, $\|x+y\| \, \|x-y\| \geqslant 2$ y, a fortiori, $\|x+y\|^2 + \|x-y\|^2 \geqslant 4$.

(iii) Si S es una curva de Radon (la B–ortogonalidad es simétrica) [7], entonces $\|x+y\| \, \|x-y\| \geqslant 2$, para todo par $x, y \in S$ tal que $x \perp y$.

(iv) Si S es un exágono regular, $\|x+y\|^2 + \|x-y\|^2 \geqslant 9/2$, para todo par $x, y \in S$ tal que $x \perp y$, (nótese que, en este caso, la B–ortogonalidad es simétrica).

(v) Si S es un octógono regular, $2\|x+y\| \, \|x-y\| \leqslant 4 \leqslant \|x+y\|^2 + \|x-y\|^2$, para todo par $x, y \in S$ tal que $x \perp y$.

(Cabe añadir que si $\approx$ representa a la ortogonalidad Isósceles ($x \perp y$ cuando $\|x+y\| = \|x-y\|$), o Pitagórica ($x \perp y$ cuando $\|x+y\|^2 = \|x\|^2 + \|y\|^2$), [9], y si S es un octógono regular, entonces se verifica (?), con $=$ en el lugar de $\sim$).

Por consiguiente, no es cierto que (?) cuando $\approx$ es la B–ortogonalidad y $\sim$ es el signo $\geqslant$, caracterice a los espacios pH, aunque puede mantenerse la <u>conjetura</u> de que sí los caracteriza cuando el signo es el $\leqslant$. (Si esto fuera cierto, entonces (i) cuando $\sim$ es $\leqslant$, cuya demostración es complicada, sería un corolario de ella).

Por otra parte, ¿cuáles son los espacios para los que $\|x+y\|\,\|x-y\|\geqslant 2$, para todo par $x,y\in S$ tal que $x\perp y$?. ¿Son aquellos para los que la B–ortogonalidad es simétrica?. (Recuérdese que si $\dim E\geqslant 3$ y si la B–ortogonalidad es simétrica, entonces E es pH).

GENERALIZACION DE CARLSSON

Además de "debilitaciones" como las antes citadas, también se han dado bastantes "generalizaciones" de la igualdad del paralelogramo (ver [1]). Una de las más notables es la de **Carlsson** [5, 6] según la cual E es pH si y solo si,

$$\boxed{\forall\, x,y\in E,\quad \sum_{k=1}^{m} a_k\|b_kx+c_ky\|^2\sim 0} \tag{6}$$

donde a_k, b_k, c_k son números reales dados tales que,

$$\boxed{a_k\neq 0;\quad b_jc_k\neq b_kc_j,\ (j\neq k);\quad \sum_{k=1}^{m} a_kb_k^2=\sum_{k=1}^{m} a_kc_k^2=\sum_{k=1}^{m} a_kb_kc_k=0.} \tag{7}$$

(Nótese que la igualdad del paralelogramo corresponde a $m=4$ y a cierta elección de los números a_k, b_k, c_k).

La primera demostración de Carlsson, del año 1962, era muy complicada y solo se refería al caso en que $\sim$ es $=$. El enunciado anterior corresponde a la demostración, mucho más simple que aquella, dada por el mismo autor en 1965. Algo más sencilla parece ser la que hace Amir ([1], pág. 13), utilizando como lema un resultado de Frechet, que dice que una función continua $g:\mathbb{R}\longrightarrow\mathbb{R}$, que verifica

$$\sum_{k=0}^{m}(-1)^{m-k}\begin{bmatrix} m\\ k\end{bmatrix} g(r+ks)=0,$$

para todo par $r,s\in\mathbb{R}$, es un polinomio de grado $<m$.

A la vista de todo lo anterior, parece lógico preguntarse si también caracteriza a los espacios pH la propiedad,

$$\boxed{\forall\, x,y\in S,\quad \sum_{k=1}^{m} a_k\|b_kx+c_ky\|^2\sim 0} \tag{??}$$

donde los números a_k, b_k, c_k están en las condiciones de (7).

Si la respuesta a esa pregunta fuera afirmativa, se obtendrían interesantes consecuencias en la pequeña parcela de la geometría de los espacios normados en la que se enmarcan estas páginas, por lo que, antes de concluirlas, merece la pena hacer algunas sencillas consideraciones sobre ella.

A tal fin, llamaremos **m–fórmula asociada a los pares (b_k,c_k)**, $(k=1,\ldots,m)$, tales que $b_jc_k-b_kc_j\neq 0$, $(j\neq k)$, a una expresión,

$$\sum_{k=1}^{m} a_k\|b_kx+c_ky\|^2\sim 0,$$

cuyos "coeficientes" a_k son $\neq 0$ y tales que,

$$\sum_{k=1}^{m} a_kb_k^2=\sum_{k=1}^{m} a_kc_k^2=\sum_{k=1}^{m} a_kb_kc_k=0. \tag{8}$$

Estas igualdades pueden verse como un sistema lineal homogéneo de tres ecuaciones en las incógnitas $a_1,\ldots,a_m$, cuyos menores de orden 3,

$$|i,j,k|=\det\begin{bmatrix} b_i^2 & c_i^2 & b_ic_i\\ b_j^2 & c_j^2 & b_jc_j\\ b_k^2 & c_k^2 & b_kc_k\end{bmatrix}=(b_ic_j-b_jc_i)(b_jc_k-b_kc_j)(b_kc_i-b_ic_k),$$

son todos distintos de cero. Por consiguiente 4 es el menor valor de m para el que existen m–fórmulas, y las m–uplas $(a_1,\ldots,a_m)$ que son solución de dicho sistema están en un subespacio $m-3$ dimensinal de $\mathbb{R}^m$. En particular, para cada elección de los pares $(b_1,c_1),\ldots,(b_4,c_4)$, existe esencialmente una única (una y sus proporcionales) 4–fórmula.

Supongamos ahora (no implica restricción alguna) que todos los b_k son $\geqslant 0$ y que los pares $(b_1,c_1),\ldots,(b_m,c_m)$ están ordenados en el sentido positivo de $\mathbb{R}^2$, es decir que $b_jc_k-b_kc_j>0$, cuando $j<k$. Entonces, $|i,j,k|>0$, siempre que $i<j<k$, y resulta fácil ver que en toda m–fórmula hay, al menos, dos a_k de cada signo. Más aún, si solo hay dos a_k de un cierto signo, entonces no tienen índices consecutivos (suponemos al decir esto que 1 sigue a m).

Sin entrar en más detalles (puramente técnicos), considérese el caso en que $a_4,\ldots,a_m$ son positivos y las siguientes igualdades que se obtienen aplicando la fórmula de Cramer,

$$\begin{aligned}
|1,2,3|\,a_1&=-|2,3,4|\,a_4-\ldots-|2,3,m|\,a_m<0,\\
|1,2,3|\,a_2&=|1,3,4|\,a_4+\ldots+|1,3,m|\,a_m>0,\\
|1,2,3|\,a_3&=-|1,2,4|\,a_4-\ldots-|1,2,m|\,a_m<0.
\end{aligned}$$

Como consecuencia de ello, si una m–fórmula tiene dos a_k de un signo y los demás del otro (cosa que ocurre, por ejemplo, con toda 5–fórmula), entonces el mismo argumento de Day [7] con las elipses de Loewner, sirve para probar que

(??) caracteriza a los espacios pH, lo cual ya fue señalado por Oman [12], aunque para el caso particular en que los pares (b_k,c_k) relativos a los dos a_k del signo singular son precisamente (1,0) y (0,1)).

Señalemos por último que toda m–fórmula se puede escribir como relación $\sim$ entre p y q–fórmulas que verifican ciertas propiedades. Por ejemplo:

Si tomamos $m-3$ vectores linealmente independientes,

$$(a_{44},\dots,a_{m4}),\dots,(a_{4m},\dots,a_{mm}),$$

con todas sus coordenadas positivas, y resolvemos para cada uno de ellos el sistema lineal homogéneo (8), obtenemos una "base",

$$(a_{14},a_{24},a_{34},a_{44},\dots,a_{m4}),\dots,(a_{1m},a_{2m},a_{3m},a_{4m},\dots,a_{mm}),$$

del espacio de coeficientes de las m–fórmulas asociadas a los pares (b_k,c_k), en la que los a_{1k} y a_{3k} son negativos, y todos los restantes a_{jk} son positivos.

Por tanto, cualquier m–upla, $(a_1,\dots,a_m)$, de coeficientes de una de dichas m–fórmulas es tal que

$$(a_1,\dots,a_m)=\lambda_4(a_{14},\dots,a_{m4})+\dots+\lambda_m(a_{1m},\dots,a_{mm}),$$

de donde se sigue que la tal m–fórmula puede escribirse

$$\sum_{\lambda_j>0}\lambda_j\sum_{k=1}^{m}a_{kj}\|b_kx+c_ky\|^2\ \sim\ -\sum_{\lambda_j<0}\lambda_j\sum_{k=1}^{m}a_{kj}\|b_kx+c_ky\|^2,$$

es decir como una relación $\sim$ entre sendas p y q–fórmulas, $(p,q\leqslant m)$, con primeros y terceros coeficientes negativos, y los restantes no negativos.

Supongamos ahora que tomamos $(a_{44},\dots,a_{m4})=(1,0,\dots,0)$ y los restantes $(a_{4j},\dots,a_{mj})$, con $a_{4j}=0$. Procediendo como antes, obtenemos que toda m–fórmula puede escribirse como relación $\sim$ entre la (única salvo proporcionales) 4–fórmula asociada a los pares (b_k,c_k), $(k=1,2,3,4)$, y una $(m-1)$–fórmula asociada a los pares (b_k,c_k), $(k=1,2,3,5,\dots,m)$.

En fin, quién sabe si consideraciones como éstas, que rompían hace años la cabeza del redactor de estas líneas, pueden resultar útiles a la hora de buscar una respuesta a la pregunta (??),....

(Mi más profundo agradecimiento al Departamento de Matemáticas de la Universidad de Los Andes (Mérida, Venezuela), cuya amable invitación en septiembre de 1993 dio lugar a la charla que recogen estas notas. Nunca había cruzado ese mar tan grande que nos une, y pocas veces me había sentido tan bien como estando allí con los buenos amigos Diomedes Bárcenas, Luisa Sánchez, Antonio Tineo y tantos otros.)

REFERENCIAS

1. D. AMIR, Characterizations of Inner Product Spaces, Birkhäuser Verlag, Basel, 1986.
2. C. BENITEZ, Una propiedad de la ortogonalidad de Birkhoff y una caracterización de espacios prehilbertianos, *Collect. Math.* 26 (1975) 211–218.
3. C. BENITEZ y M. del RIO, Characterization of inner product spaces through rectangle and square inequalities, *Rev. Roum. Math. Pures Appl.* 29 (1984) 543–546.
4. G. BIRKHOFF, Orthogonality in linear metric spaces, *Duke Math. J.* 1 (1935) 169–172.
5. S.O. CARLSSON, Orthogonality in normed linear spaces, *Arkiv für Matem.* 4 (1962) 297–318.
6. S.O. CARLSSON, A characteristic property of Euclidean spaces, *Arkiv für Matem.* 5 (1965) 327–330.
7. M.M. DAY, Some characterizations of inner product spaces, *Trans. Amer. Math. Soc.* 62 (1947) 320–337.
8. M.M. DAY, Some criteria of Kasahara and Blumenthal for inner product spaces, *Proc. Amer. Math. Soc.* 10 (1959) 92–100.
9. R.C. JAMES, Orthogonality in normed linear spaces, *Duke Math. J.* 12 (1945) 291–301.
10. R.C. JAMES, Orthogonality and linear functionals in normed linear spaces, *Trans. Amer. Math. Soc.* 61 (1947) 265–292.
11. P. JORDAN y J. von NEUMANN, On inner products in normed linear spaces, *Ann. of Math.* 36 (1935) 705–718.
12. J. OMAN, Norm identities which characterize inner product spaces, in "The Geometry of Metric and Linear Spaces", Michigan 1974, Springer Lecture notes 490, 116–133.
13. I.J. SCHOENBERG, A remark on M.M. Day's characterization of Hilbert Space, *Proc. Amer. Math. Soc.* 3 (1952) 961–964.

DE LA DESIGUALDAD DE KHINTCHINE A LAS DESIGUALDADES DEL BUEN λ

WILFREDO URBINA

Departamento de Matemáticas, Facultad de Ciencias,

Universidad Central de Venezuela.

Mérida, 29 de Julio de 1993.-

De la desigualdad de Khintchine a las desigualdades del buen λ

Wilfredo Urbina
Dpto. de Matemáticas.Facultad de Ciencias. Universidad Central de Venezuela.Apartado 47195 Los Chaguaramos. Caracas 1041-A.

§1. Introducción

Las funciones de Rademacher han sido objeto de intensa investigación desde que Hans Rademacher inicio su estudio en 1922 [15]. Dichas funciones combinan en su naturaleza tanto una característica analítica como una probabilística y precisamente en su estudio hasta hoy continúan vigentes estas dos vertientes y es fuente inagotable de inspiración.

En el presente trabajo queremos mostrar como una desigualdad que ellas verifican, la desigualdad de Khintchine, ha servido de fuente de inspiración para la obtención de importantes desigualdades para martingalas y de técnicas desarrolladas para la obtención de ellas. Nos referimos especificamente a las desigualdades de Burkholder-Davis-Gundy y de las desigualdades del buen λ.

§2 . La desigualdad de Khintchine

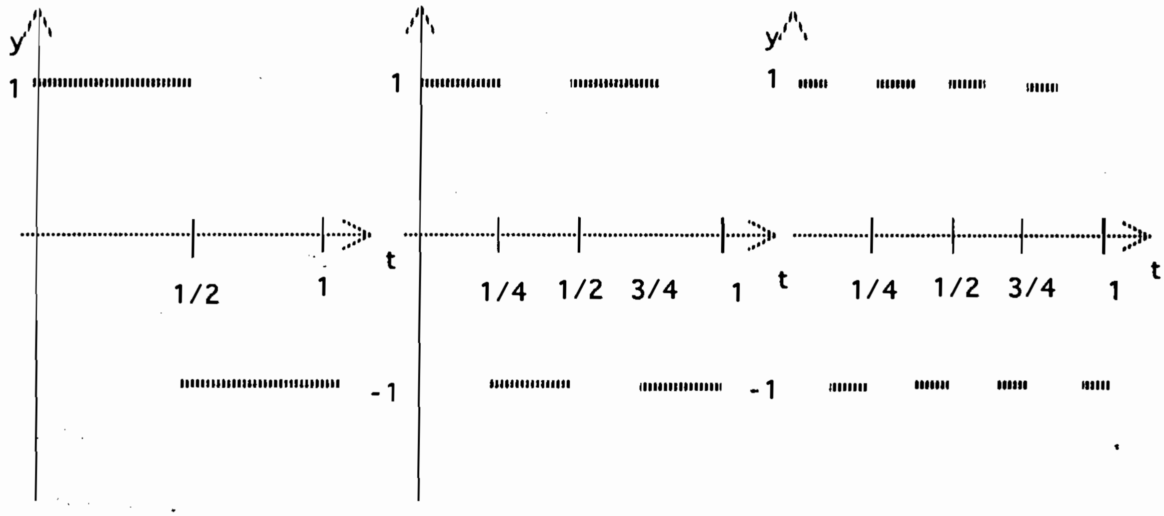
Como es usual, consideremos el intervalo $[0,1)$, $\mathfrak{B}[0,1)$ la familia de los conjuntos Borelianos en $[0,1)$, es decir,la σ-álgebra de los conjuntos generados por los abiertos de $[0,1)$ y m la medida de Lebesgue, denotemos a estas escogencias por el triplete por $([0,1), \mathfrak{B}[0,1), m)$.

Las funciones de Rademacher se pueden definir entonces como:

$$r_n(t) = \text{signo}\,(\text{sen}\,(2^n \pi t)) \qquad (1)$$

ó equivalentemente como $r_1(t) = \text{signo}\,(\text{sen}\,(2\pi t))$ y $r_n(t) = r_1(2^{n-1}t)$ para $n > 1$, (obsérvese que sen $(2^n \pi x)$ cambia de signo en intervalos $I_n^k = \left[\frac{k-1}{2^n}, \frac{k}{2^n} \right)$).

Las gráficas de las tres primeras funciones de Rademacher son las siguientes:



r_1 r_2 r_3

Obsérvese que para todo n

$$\int_0^1 r_n(t)\,dt = 0 \quad y \quad \int_0^1 (r_n(t))^2\,dt = 1.$$

Además, es fácil ver que, las funciones de Rademacher son estocásticamente independientes, es decir, si $\delta_1,\ldots,\delta_n$ es una sucesión de + 1 y -1, entonces

$$m\{t\in[0,1): r_1(t)=\delta_1,\ldots,r_n(t)=\delta_n\} = m\{t\in[0,1): r_1(t) = \delta_1\}\ldots\ m\{t\in[0,1): r_n(t) = \delta_n\}.$$

Esta propiedad implica la identidad

$$\int_0^1 r_{n_1}^{\alpha_1}(t)\ r_{n_2}^{\alpha_2}(t)\ldots\ r_{n_k}^{\alpha_k}(t)\,dt = \int_0^1 r_{n_1}^{\alpha_1}(t)\,dt\ldots\int_0^1 r_{n_k}^{\alpha_k}(t)\,dt, \qquad (2)$$

para $\alpha_1,\ldots,\alpha_k \in \mathbb{N}$ y más generalmente si $f_1,\ldots, f_k$ son funciones integrables en [0,1) tenemos

$$\int_0^1 f_1(r_{n_1}(t))\ldots\ f_k(r_{n_k}(t))\ dt = \int_0^1 f_1(r_{n_1}(t))\,dt\ldots\int_0^1 f_k(r_{n_k}(t))\,dt. \qquad (2')$$

En particulas éstas identidades implican que $\{r_n\}$ es un sistema ortonormal en $L^2([0,1))$ aunque es fácil de ver que no es completo.

Se puede probar que si $\{a_n\} \in \ell_2$, entonces la serie $\sum_{n=1}^{\infty} a_n r_n(t)$ converge c.s. (esto es un teorema de Rademacher, Kolmogorov y Khintchine). La desigualdad de Khintchine, que como hemos dicho es de gran importancia, da más información sobre éstas series

Teorema (Desigualdad de Khintchine): Sea $\{a_n\} \in \ell_2$ entonces la serie $\sum_{n=1}^{\infty} a_n r_n(t)$ pertenece a L^p para todo $p > 0$ y más precisamente, existen constantes positivas A_p, B_p, que dependen sólo de p, tal que para todo $\{a_n\} \in \ell_2$

$$A_p \left(\sum_{n=1}^{\infty} |a_n|^2 \right)^{1/2} \leq \left(\int_0^1 | \sum_{n=1}^{\infty} a_n r_n(t)|^p dt \right)^{1/p} \leq B_p \left(\sum_{n=1}^{\infty} |a_n|^2 \right)^{1/2}, \qquad (3)$$

en particular, para todo $N \in \mathbf{N}$

$$A_p \left(\sum_{n=1}^{N} |a_n|^2 \right)^{1/2} \leq \left(\int_0^1 | \sum_{n=1}^{N} a_n r_n(t)|^p dt \right)^{1/p} \leq B_p \left(\sum_{n=1}^{N} |a_n|^2 \right)^{1/2}. \qquad (3')$$

Observación: La desigualdad de Khintchine da una prueba indirecta que las funciones de Rademacher no son un sistema completo ya que en el espacio generado por las funciones de Rademancher todas las métricas L^p son equivalentes. Obsérvese además que por ésta desigualdad el subespacio cerrado en L^p generado por las funciones de Rademacher es isomorfo a ℓ_2.

Demostración

Por su importancia daremos dos pruebas diferentes de ésta desigualdad.

I) Probemos primero la desigualdad derecha de (3). Obsérvese que

$$\int_0^1 e^{\xi a_n r_n(t)} dt = \sum_{k=1}^{2^n} \int_{(k-1)/2^n}^{k/2^n} \exp [\xi\, a_n\, r_n(t)]\, dt = \frac{1}{2^n} \sum_{k\ par} \exp[\xi a_n] + \frac{1}{2^n} \sum_{k\ impar} \exp[-\xi a_n]$$

$$= \frac{e^{\xi a_n} + e^{-\xi a_n}}{2} = \cosh (\xi\, a_n).$$

Entonces por independencia obtenemos, para cada N

$$\int_0^1 \exp\left[\xi \sum_{n=1}^{N} a_n r_n(t)\right] dt = \prod_{n=1}^{N} \int_0^1 \exp(\xi\, a_n r_n(t))\, dt = \prod_{n=1}^{N} \cosh(\xi\, a_n)$$

$$\leq \exp\left(\xi^2 \sum_{n=1}^{N} a_n^2\right),$$

ya que $\cosh x \leq e^{x^2}$.

Ahora bien, como $\{a_n\} \in l_2$, hemos dicho que $\sum_{n=1}^{\infty} a_n r_n(t)$ converge c.s.,

aplicando entonces el Lema de Fatou, obtenemos

$$\int_0^1 \exp\left[\xi \sum_{n=1}^{\infty} a_n r_n(t)\right] dt \leq \exp\left(\xi^2 \sum_{n=1}^{\infty} a_n^2\right).$$

Si ahora normalizamos tomando $\sum_{n=1}^{\infty} a_n^2 = 1$, entonces como

$$\exp\left(\xi \left| \sum_{n=1}^{\infty} a_n r_n(t) \right|\right) \leq \exp\left(\xi\left[\sum_{n=1}^{\infty} a_n r_n(t)\right]\right) + \exp\left(-\xi\left[\sum_{n=1}^{\infty} a_n r_n(t)\right]\right)$$

tenemos

$$\int_0^1 \exp\left(\xi \left| \sum_{n=1}^{\infty} a_n r_n(t) \right|\right) dt \leq 2\, e^{\xi^2}.$$

Sea ahora $\lambda(\xi) = m\{ t \in [0,1) : |\sum_{n=1}^{\infty} a_n r_n(t)| > \xi \}$, la función de distribución de $\sum_{n=1}^{\infty} a_n r_n(t)$, entonces usando la desigualdad de Tchebychev (con la función $e^{(\xi/2)x}$)

tenemos

$$\lambda(\xi) = m\{ t \in [0,1) : |\sum_{n=1}^{\infty} a_n r_n(t)| > \xi \} \leq e^{-(\xi/2)\xi} \int_0^1 \exp\left(\frac{\xi}{2} \left|\sum_{n=1}^{\infty} a_n r_n(t)\right|\right) dt$$

$$\leq e^{-\xi^2/2}\, 2\, e^{(\xi/2)^2} = 2\, e^{-\xi^2/4}.$$

Teniendo esta estimación de la función de distribución, consideremos

$$\left(\int_0^1 \left| \sum_{n=1}^{\infty} a_n r_n(t) \right|^p dt \right)^{1/p} = \left(p \int_0^1 \left(\int_0^{|\sum_n a_n r_n(t)|} \xi^{p-1}\, d\xi \right) dt \right)^{1/p}$$

$$= \left(p \int_0^\infty \int_0^1 \chi_{\{\xi : 0 \le \xi < |\sum_n a_n r_n(t)|\}}(\xi) \xi^{p-1}\, d\xi\, dt \right)^{1/p}$$

$$\le \left(p \int_0^\infty \left[\int_0^1 \chi_{\{t \in [0,1[: |\sum_n a_n r_n(t)| > \xi\}}(t)\, dt \right] \xi^{p-1}\, d\xi \right)^{1/p}$$

$$= \left(p \int_0^\infty \xi^{p-1} \lambda(\xi) d\xi \right)^{1/p} \le \left(2p \int_0^\infty \xi^{p-1} e^{-\xi^2/4} d\xi \right)^{1/p}$$

$$= (2^p\, p\, \Gamma(\tfrac{p}{2}))^{1/p} = 2\, p^{1/p} (\Gamma(\tfrac{p}{2}))^{1/p} = B_p$$

y luego, en general, tenemos

$$\left(\int_0^1 | \sum_{n=1}^\infty a_n r_n(t)|^p\, dt \right)^{1/p} \le B_p \left(\sum_{n=1}^\infty |a_n|^2 \right)^{1/2} < \infty.$$

La desigualdad izquierda es consecuencia de ésta. De hecho, por un argumento de dualidad, basta considerar el caso $0 < p < 2$. Sea θ definido como $\theta = (2 - p/2)^{-1}$ entonces $1/2 < \theta < 1$ y $p\theta + 4(1 - \theta) = 2$, aplicando ahora la desigualdad de Hölder

$$\left(\int_0^1 \left(\sum_{n=1}^\infty a_n r_n(t) \right)^2 dt \right)^{1/2} \le \left(\int_0^1 | \sum_{n=1}^\infty a_n r_n(t)|^p\, dt \right)^{\theta/2} \left(\int_0^1 | \sum_{n=1}^\infty a_n r_n(t)|^4\, dt \right)^{(1-\theta)/2}.$$

Por la desigualdad derecha, para $p = 4$, y la ortogonalidad, tenemos

$$\left(\int_0^1 | \sum_{n=1}^\infty a_n r_n(t)|^4\, dt \right)^{(1-\theta)/2} = \left(\int_0^1 | \sum_{n=1}^\infty a_n r_n(t)|^4\, dt \right)^{2(1-\theta)/4}$$

$$\le B_4^{2(1-\theta)} \left(\sum_{n=1}^\infty |a_n|^2\, dt \right)^{(1-\theta)} = B_4^{2(1-\theta)} \left(\int_0^1 \left(\sum_{n=1}^\infty a_n r_n(t) \right)^2 dt \right)^{(1-\theta)},$$

Luego

$$\left(\int_0^1 | \sum_{n=1}^\infty a_n r_n(t) |^2\, dt \right)^{p\theta/4} \le B_4^{2(1-\theta)} \left(\int_0^1 | \sum_{n=1}^\infty a_n r_n(t) |^p\, dt \right)^{\theta/2}$$

por tanto,

$$\left(\int_0^1 | \sum_{n=1}^\infty a_n r_n(t) |^2\, dt \right)^{1/2} \le B_4^{4(1-\theta)/p\theta} \left(\int_0^1 | \sum_{n=1}^\infty a_n r_n(t) |^p\, dt \right)^{1/p}$$

y, de nuevo, por la ortogonalidad

$$B_4^{-4/p}\left(\sum_{n=1}^{\infty}|a_n|^2\right)^{1/2} \le \left(\int_0^1\left|\sum_{n=1}^{\infty} a_n\, r_n(t)\right|^p dt\right)^{1/p}. \blacksquare$$

II) Para $p = 2$ tenemos igualdad por la ortogonalidad de los funciones $\{r_n\}$.

Basta probar la desigualdad derecha de (3) para $p > 2$ ya que si $0 < \alpha < 1$ es tal que $\frac{1}{2} = \frac{1-\alpha}{p} + \frac{4}{\alpha}$, usando la desigualdad de Holder, obtenemos

$$\left(\sum_{n=1}^{\infty}|a_n|^2\right)^{1/2} = \left(\int_0^1\left|\sum_{n=1}^{\infty} a_n\, r_n(t)\right|^2 dt\right)^{1/2}$$

$$\le \left(\int_0^1\left|\sum_{n=1}^{\infty} a_n\, r_n(t)\right|^p dt\right)^{(1-\alpha)/p} \left(\int_0^1\left|\sum_{n=1}^{\infty} a_n\, r_n(t)\right|^4 dt\right)^{\alpha/4}$$

$$\le \left(\int_0^1\left|\sum_{n=1}^{\infty} a_n\, r_n(t)\right|^p dt\right)^{(1-\alpha)/p} \left(B_4\left(\sum_{n=1}^{\infty}|a_n|^2\right)^{1/2}\right)^{\alpha}$$

y como todas las normas son finitas, obtenemos la desigualdad izquierda con $A_p = B_4^{-\alpha/1-\alpha}$. Asi pues, es suficiente probar la desigualdad derecha de (6) para $p > 2$.

Más aún, dado que la norma p es una función creciente en p, basta considerar p entero.

Supongamos además que los a_n son reales y tal que $\sum_{n=1}^{\infty}|a_n|^2 = 1$. Bajo éstas hipótesis, usando la desigualdad

$$|x|^p \le p!\, e^{|x|} \le p!\,(e^x + e^{-x}), \; x \in \mathbf{R},$$

la independencia y la simetría de las $\{r_n\}$, tenemos

$$\int_0^1\left|\sum_{n=1}^{\infty} a_n\, r_n(t)\right|^p dt \le p! \int_0^1 \left[\exp\left(\sum_{n=1}^{\infty} a_n\, r_n(t)\right) + \exp\left(-\sum_{n=1}^{\infty} a_n\, r_n(t)\right)\right] dt$$

$$\le 2\, p! \prod_{n=1}^{\infty} \int_0^1 \exp(a_n\, r_n(t))\, dt,$$

y como, por un cálculo ya hecho, tenemos que

$$\int_0^1 e^{\xi a_n r_n(t)}\, dt = \cosh(\xi a_n),$$

entonces, dado que $\cosh x \le e^{x^2}$ obtenemos

$$\int_0^1 \left| \sum_{n=1}^{\infty} a_n r_n(t) \right|^p dt \le 2\,p! \prod_{n=1}^{\infty} \cosh(a_n)$$

$$\le 2\,p! \prod_{n=1}^{\infty} \exp(a_n^2) = 2\,e\,p!\ .\blacksquare$$

La desigualdad de Khintchine, tambien se puede probar de manera inmediata utilizando sistemas Gaussianos, pero ésta demostración se sale del contexto donde estamos trabajando.

La estimación que hemos obtenido para B_p es del orden $O(p)$ si $p \to \infty$. Siendo más cuidadosos podemos probar que B_p es del orden $O(p^{1/2})$ si $p \to \infty$. Constantes optimales para la desigualdad de Khintchine fueron determinadas por S.J. Szareck [18] y son las siguientes:

$A_p = 1$ si $p \ge 2$, $B_p = 1$ si $1 \le p \le 2$, $B_{2m} = ((2m)!!)^{1/2m}$, $m = 1,2,\ldots$ y $A_1 = \dfrac{1}{\sqrt{2}}$.

Obsérvese que como consecuencia de la desigualdad de Khintchine tenemos que para cualquier $p > 0$, dada una serie de Rademacher $\sum_{n=1}^{\infty} a_n r_n(t)$ y consideramos <u>cualquier</u> cambio de signo de los coeficientes a_n, es decir, considerando $\varepsilon_n \in \{-1,1\}$, $\sum_{n=1}^{\infty} \varepsilon_n a_n r_n(t)$, entonces para todo N

$$A_p \left(\sum_{n=1}^{N} |a_n|^2 \right)^{1/2} \le \left(\int_0^1 \left| \sum_{n=1}^{N} \varepsilon_n a_n r_n(t) \right|^p dt \right)^{1/p} \le B_p \left(\sum_{n=1}^{N} |a_n|^2 \right)^{1/2}$$

o dicho de otro modo

$$\left\| \sum_{n=1}^{N} \varepsilon_n a_n r_n \right\|_p \approx \left(\sum_{n=1}^{N} |a_n|^2 \right)^{1/2}$$

y como (3') se cumple podemos concluir entonces

Proposición : Si $\{a_n\} \in l_2$ entonces para cualquier $p > 0$ y para cualquier cambio de signos $\{\varepsilon_n\}$, con $\varepsilon_n \in \{-1,1\}$, tenemos que, para todo N

$$\| \sum_{n=1}^{N} \varepsilon_n a_n r_n \|_p \approx \| \sum_{n=1}^{N} a_n r_n \|_p . \tag{4}$$

En particular, existe C_p tal que

$$\| \sum_{n=1}^{N} \varepsilon_n a_n r_n \|_p \leq C_p \| \sum_{n=1}^{N} a_n r_n \|_p . \tag{4'}$$

§3. Las desigualdades de Paley

Por medio de las funciones de Rademacher podemos definir dos sistemas ortonormales completos en $L^2([0,1),m)$. En primer lugar,las funciones de Walsh están definidas de la siguiente manera. Sea

$$\psi_0(t) = 1, \text{ para } t \in [0,1) ,$$

y para $n \geq 1$ si $2n = 2^{n_1} + 2^{n_2} + ... + 2^{n_k}$, con $0 < n_1 < n_2 < ... < n_k$, entonces

$$\psi_n(t) = r_{n_1}(t)\, r_{n_2}(t) \,..\, r_{n_k}(t) \quad n \geq 1. \tag{5}$$

Obsérvese que el sistema de Walsh contiene al sistema de Rademacher, ya que

$$\psi_{2^{n-1}} = r_n \tag{5'}$$

El sistema de Walsh es una completación de las funciones de Rademacher, ya que intuitivamente se agregan todas las funciones ortogonales a ellas, formalmente R.E.A.C. Paley [14] probó que el sistema de Walsh es un sistema ortonormal completo en $L^2\,[0,1)$.

Cualquier función $f \in L^1([0,1),m)$ puede ser desarrollada (formalmente) por medio de las funciones de Walsh:

$$f(t) \sim \sum_{n=1}^{\infty} c_n \psi_n(t), \text{ donde } c_n = \int_0^1 f(t)\, \psi_n(t)\, dt , \tag{6}$$

que se llama expansion de Walsh-Fourier de f, denotada por W(f) y denotaremos la suma parcial N-ésima $\sum_{n=0}^{N-1} c_n \psi_n(t)$ por $W_N(f)$.

En segundo lugar, las funciones de Haar, el otro sistema derivado del las funciones de Rademacher, se pueden definir, de las siguientes maneras

$$\chi_k^{(n)}(t) = \sqrt{2^{n-1}}\, (\chi_{[(2k-2)/2^n, (2k-1)/2^n)} - \chi_{[(2k-1)/2^n, k/2^{n-1})}) \tag{7}$$

$$= \sqrt{2^{n-1}}\,\chi_1^{(1)}(2^{n-1}t-(k-1)) = r(2^{n-1})\, r_1(2^{n-1}t-(k-1)),\ k=1,\ldots,2^{n-1} \quad (7')$$

Obsérvese que, por (7') $\chi_k^{(n)}$ se obtienen por dilataciones y traslaciones de r_1, y que de manera inmediata, por la definición de las funciones de Haar, se tiene

$$r_n(t) = \frac{1}{\sqrt{2^{n-1}}} \sum_{k=1}^{2^{n-1}} \chi_k^{(n)} .$$

La familia $\{\chi_n^{(k)}\}$ es tambien un sistema ortonormal completo en $L^2([0,1))$.

Además, existe une estrecha relación entre las funciones de Walsh y las de Haar. De hecho tenemos que las funciones de Walsh $\{\psi_k^{(n)}\}$ se puede expresar como combinaciones lineales de las funciones de Haar $\{\chi_k^{(n)}\}$ y vice versa (véase S. Kaczmarz y H. Steinhaus [11] y Schipp, Wade y Simon [16]).

Ahora, dado $f \in L^1([0,1),m)$ la expansión de f por medio del sistema de Haar, que llamaremos expansión de Haar-Fourier de f, está dado por

$$d_0 + \sum_{n=1}^{\infty} \sum_{k=1}^{2^{n-1}} d_k^{(n)} \chi_k^{(n)}(t) \text{ donde } d_k^{(n)} = \int_0^1 f(t)\, \chi_k^{(n)}(t)\, dt. \quad (8)$$

Por comodidad la expansión de Haar-Fourier de f la denotaremos simplemente como $\sum_{n,k} d_k^{(n)} \chi_k^{(n)}(t)$, aunque debe entenderse la variación de los índices como en la expresión previa. La expansión de Haar-Fourier de una función f se denota como Hf y sus sumas parciales por $H_N f$, que se define como $H_N f(t) = \sum_{n=0}^{N-1} \sum_{k=1}^{2^{n-1}} d_k^{(n)} \chi_k^{(n)}(t)$.

Walsh probó que las 2^N-esima sumas parciales de las series de Walsh y de Haar $W_{2^N} f$ y $H_{2^N} f$ son iguales para todo N, mediante la relación que existe entre las funciones de Walsh y las de Haar ya mencionada.

Finalmente, asi como para las funciones de Rademacher tenemos la desigualdad de Khinchine, para las funciones de Walsh y de Haar, que son derivadas de ellas, existen también desigualdades análogas. Consideremos el caso de las funciones de Haar, el caso de las funciones de Walsh es completamente análogo.

Sea f una función integrable en [0,1), sea y $\sum_{n,k} d_k^{(n)} \chi_k^{(n)}(t)$ su expansión de Haar-Fourier, definimos

$$S(f) = \left[\sum_{n,k} (d_k^{(n)} \chi_k^{(n)}(t))^2 \right]^{1/2} \tag{9}$$

la **función cuadrática** asociada a f. Entonces, R.E.A.C Paley [14], probó que existen constantes c_p y C_p (independientes de f) tales que

$$c_p \| S(f) \|_p \leq \| f \|_p \leq C_p \| S(f) \|_p , \quad 1 < p < \infty . \tag{10}$$

En realidad Paley probó ésta desigualdad para series de Walsh. La versión para series de Haar fue dada por Marcinkiewicz [12].

El papel de la función cuadrática en esta desigualdad es clave y como veremos esta noción es de gran importancia tambien en contextos más generales tanto en Probabilidad como en Análisis. Para un interesante recuento sobre el desarrollo de éste concepto véase Stein [17].

Otra desigualdad importante es la desigualdad maximal de Hardy-Littlewood : existe una constante c_p (independiente de f) tal que

$$c_p \| f^* \|_p \leq \| f \|_p \leq \| f^* \|_p , \quad 1 < p < \infty , \tag{11}$$

donde $f^* = \sup_N | \sum_{n,k}^{N} d_k^{(n)} \chi_k^{(n)} |$, es la función maximal de las sumas parciales.

De las dos desigualdades anteriores se obtiene entonces la siguiente desigualdad: existen constantes c_p y C_p (independientes de f) tal que

$$c_p \| S(f) \|_p \leq \| f^* \|_p \leq C_p \| S(f) \|_p , \quad 1 < p < \infty . \tag{12}$$

La desigualdad, análoga a (4'), obtenida para éste caso tambien por R.E.A.C Paley [14], es la siguiente: dados $d_k^{(n)} \in \mathbf{R}$, $\varepsilon_k^{(n)} \in \{-1,1\}$ entonces existe una constante C_p (independiente de los $d_n^{(k)}$ y de los $\varepsilon_k^{(n)}$) tal que

$$\| \sum_{n,k} \varepsilon_k^{(n)} d_k^{(n)} \chi_k^{(n)} \|_p \leq C_p \| \sum_{n,k} d_k^{(n)} \chi_k^{(n)} \|_p , \quad 1 < p < \infty . \tag{13}$$

Es de notar que, debido precisamente a la desigualdad de Khintchine, las dos desigualdades de Paley son equivalentes.

Como veremos más adelante estas desigualdades son consecuencia inmediata de desigualdades mucho más generales para martingalas. Sin embargo muchas de las ideas en ese contexto general estan inspiradas por ellas e incluso , como lo notó Richard Gundy (véase comentario en [3]), algunas de las pruebas originales se pueden adaptar al ese caso general.

§4. Breve introducción a la Teoría de Martingalas

Las generalizaciones de los sistemas ortonormales estudiados, en el campo de las probabilidades, son las martingalas.

Sea $(\Omega,\mathscr{F},\mathbf{P})$ un espacio de probabilidad y consideremos una filtración $\{\mathscr{F}_n\}$ en Ω (es decir, $\{\mathscr{F}_n\}$ es una una sucesión creciente, $\mathscr{F}_{n+1}\supseteq\mathscr{F}_n$ de σ-álgebras contenidas en la σ- álgebra $\mathscr{F}$)

Definición: Un proceso $X = \{X_n\}_n$ es una $\{\mathscr{F}_n\}$-**martingala** ($\{\mathscr{F}_n\}$-supermartingala, $\{\mathscr{F}_n\}$-submartingala) si:

i) Para cada n, $X_n \in L^1(\Omega)$, es decir, si $\mathbf{E}[|X_n|] < \infty$.

ii) $\{ X_n \}$ es $\{\mathscr{F}_n\}$-adaptado, es decir, para cada n X_n es $\mathscr{F}_n$- medible.

iii) Para todo $m,n \in \mathbf{N}$, si $m > n$, entonces

$$\mathbf{E}[X_m | \mathscr{F}_n] = X_n \quad \text{c.s.} \qquad (\leq, \geq \text{ respectivamente}). \tag{14}$$

Observaciones:

1) La condición (14) es la propiedad clave de la definición de estos procesos y ella es equivalente, por la definición de esperanza condicional, a

$$\forall A \in \mathscr{F}_n \qquad \int_A X_m \, d\mathbf{P} = \int_A X_n \, d\mathbf{P} \quad (\leq, \geq \text{ respectivamente}). \tag{15}$$

(15) a su vez equivale a pedir que para toda variable aleatoria Y $\mathscr{F}_n$ - medible

$$\int_\Omega Y X_m \, d\mathbf{P} = \int_\Omega Y X_n \, d\mathbf{P} \quad (\leq, \geq \text{ respectivamente}). \tag{16}$$

2) La condición iii) es equivalente, por inducción, a

iii)' Para todo $n\in\mathbf{N}$, $\mathbf{E}[X_{n+1}|\mathcal{F}_n] = X_n$ c.s.($\leq$, $\geq$ respectivamente). (17)

3) Finalmente, si definimos $\Delta X_1 = X_1$ y si $n>1$, $\Delta X_n = X_n - X_{n-1}$ (o lo que es igual definir $\Delta X_n = X_n - X_{n-1}$ para todo n suponiendo que $X_0 = 0$) y llamamos $\Delta X = \{\Delta X_n\}$ a la sucesión de diferencias del proceso X entonces iii) tambien equivale a pedir la siguiente condición

iii)" Para todo $n\in\mathbf{N}$, $\mathbf{E}[\Delta X_n | \mathcal{F}_n] = 0$ c.s. ($\leq$, $\geq$ respectivamente). (14')

La condición iii)" a veces se expresa en forma "analítica" de la siguiente manera

$$\int_\Omega \Psi(X_1,\ldots,X_{n-1})\, \Delta X_n\, d\mathbf{P} = 0 \quad (\leq, \geq \text{ respectivamente}).$$

para toda funcion Ψ: $\mathbf{R}^{n-1} \to \mathbf{R}$, medible y acotada.

En el caso de martingalas, (14') se conoce como la propiedad de **ortogonalidad de los incrementos.** Es claro que

$$X_n = \sum_{k=1}^{n} \Delta X_k \tag{18}$$

y por tanto las martingalas son sumas de variables aleatorias ortogonales en el sentido de (14). Ademas, es claro que es igual conocer una martingala X que su sucesion de incrementos ΔX.

Ejemplos de martingalas

1) El típico ejemplo de una martingala, y del cual las martingalas son una directa generalización, son las sumas de variables aleatorias independientes $\{\zeta_n\}$ centradas, $\mathbf{E}[\zeta_n] = 0$. En ese caso las condiciones i) y ii) son inmediatas y verificaremos iii)' :

Por la observación anterior $\mathcal{F}_n = \sigma(\zeta_1, \zeta_1 + \zeta_2, \ldots, \sum_{k=1}^{n}\zeta_k) = \sigma(\zeta_1, \ldots, \zeta_n)$ y por la independencia de los ζ_n :

$$\mathbf{E}[\sum_{k=1}^{n+1}\zeta_k | \mathcal{F}_n] = \mathbf{E}[\sum_{k=1}^{n}\zeta_k + \zeta_{n+1}|\mathcal{F}_n] = \mathbf{E}[\sum_{k=1}^{n}\zeta_k|\mathcal{F}_n] + \mathbf{E}[\zeta_{n+1}|\mathcal{F}_n]$$

$$= \sum_{k=1}^{n}\zeta_n + \mathbf{E}[\zeta_{n+1}] = \sum_{k=1}^{n}\zeta_n .$$

En particular, el paseo al azar simétrico y las sumas parciales de series de Rademacher, $\sum_{k=1}^{\infty} a_k r_k(t)$, son martingalas .

Sumas de variables aleatorias independientes $\{\zeta_n\}$ tales que $\mathbf{E}(\zeta_n) \geq 0$ ó que $\mathbf{E}(\zeta_n) \leq 0$ son ejemplos de submartingalas ó supermartingalas, respectivamente.

2) Dada $X \in L^1(\Omega)$, si definimos $X_n = \mathbf{E}[X | \mathcal{F}_n]$, entonces $\{X_n\}$ es una $\{\mathcal{F}_n\}$-martingala. La condición i) se obtiene de forma inmediata ya que

$$\mathbf{E}[\,|X_n|\,] = \mathbf{E}[\,|\mathbf{E}[X | \mathcal{F}_n]|\,] \leq \mathbf{E}[\,\mathbf{E}[\,|X|\,|\mathcal{F}_n]\,] = \mathbf{E}[|X|],$$

ii) es consecuencia de la definición de esperanza condicional y por las propiedades de la esperanza condicional, tenemos :

$$\mathbf{E}[X_m | \mathcal{F}_n] = \mathbf{E}[\,\mathbf{E}[X | \mathcal{F}_m] | \mathcal{F}_n] = \mathbf{E}[X | \mathcal{F}_n] = X_n .$$

Como veremos en el desarrollo de la teoria, muchas de las martingalas interesantes son de este tipo.

3) Sea $W_n = \sum_{k=1}^{n-1} c_k \psi_k(t)$, donde $\{c_k\}$ es una sucesión numérica (real ó compleja) y $\{\psi_k\}$ son las funciones de Walsh . Entonces, como notó R. Gundy [6] ,$\{W_{2^n}\}$ es una martingala ya que

$$W_{2^n} = W_{2^{n-1}} + \sum_{k=2^{n-1}}^{2^n-1} c_k \psi_k(t) = W_{2^{n-1}} + r_n(t) \sum_{k=0}^{2^n-1} c_{2^{n-1}+k} \psi_k(t)$$

Luego,

$$\mathbf{E}[W_{2^n} | \sigma(\psi_k : k \leq 2^{n-1})] = \mathbf{E}[W_{2^n} | \sigma(r_k : k \leq n)] = W_{2^{n-1}} ,$$

por la independencia de las funciones de Rademacher.

Análogamente si $H_N = c_0 + \sum_{n=1}^{N} \sum_{k=1}^{2^{n-1}} c_k^{(n)} \chi_k^{(n)}(t)$, donde $\{c_k^{(n)}\}$ es una sucesión numérica (real ó compleja) y $\{\chi_k^{(n)}\}$ son las funciones de Haar, se puede probar, directamente o utilizando la observación de Walsh mencionada anteriormente que W_{2^n} coincide con H_{2^n}, que $\{H_{2^n}\}$ es una martingala.

Definición : Dado un proceso $Y = \{Y_n\}_n$ $\{\mathcal{F}_n\}$-predecible (es decir, Y_n es $\mathcal{F}_{n-1}$ medible para cada n) uniformemente acotado, entonces dada $X = \{X_n\}_n$ una $\{\mathcal{F}_n\}$ -martingala definimos el proceso estocástico $(Y\cdot X) = \{(Y\cdot X)_n\}_n$, que llamaremos **transformada de X por Y** como

$$(Y\cdot X)_n = \sum_{k=1}^{n} Y_n \Delta X_k = \sum_{k=1}^{n} Y_n (X_k - X_{k-1}) \text{ para todo n.} \qquad (19)$$

Obsérvese que $(Y\cdot X)_n = (Y\cdot X)_{n-1} + Y_n \Delta X_n$ y por tanto

$$\Delta(Y\cdot X)_n = (Y\cdot X)_n - (Y\cdot X)_{n-1} = Y_n \Delta X_n .$$

Proposición : Sea $X = \{X_n\}_n$ una $\{\mathcal{F}_n\}$ -martingala y sea $Y = \{Y_n\}_n$ un proceso $\{\mathcal{F}_n\}$-predecible y uniformemente acotado, la transformada de X por Y $(Y\cdot X) = \{ (Y\cdot X)_n\}_n$ es una $\{\mathcal{F}_n\}$ -martingala que llamaremos la **martingala transformada de X por Y.**

Demostración

Como $\{X_n\}_n$ es una martingala y como $\{Y_n\}_n$ es acotada uniformemente entonces para todo n, $\mathbf{E}| (Y\cdot X)_n | \leq M\, \mathbf{E}| X_n | < \infty$ y claramente $(Y\cdot X)_n$ es $\mathcal{F}_n$ - medible ya que $\{Y_n\}$ es predecible, finalmente si es $m \geq n$, por la propiedad de martingala de X

$$\mathbf{E}\,[(Y\cdot X)_m | \mathcal{F}_n] = \mathbf{E}\,[\sum_{k=1}^{m} Y_k \Delta X_k \mid \mathcal{F}_n] = \sum_{k=1}^{n} \mathbf{E}\,[Y_k \Delta X_k \mid \mathcal{F}_n] + \sum_{k=n+1}^{m} \mathbf{E}\,[Y_k \Delta X_k \mid \mathcal{F}_n]$$

$$= \sum_{k=1}^{n} Y_k (X_k - X_{k-1}) + \sum_{k=n+1}^{m} \mathbf{E}\,[\mathbf{E}\,[Y_k \Delta X_k \mid \mathcal{F}_{k-1}] \mid \mathcal{F}_n]$$

$$= \sum_{k=1}^{n} Y_k (X_k - X_{k-1}) + \sum_{k=n+1}^{m} \mathbf{E}\,[Y_k\, \mathbf{E}[\Delta X_k | \mathcal{F}_{k-1}] \mid \mathcal{F}_n] = \sum_{k=1}^{n} Y_k \Delta X_k = (Y\cdot X)_n. \blacksquare$$

Definición : Llamaremos una **sucesion multiplicadora** a todo proceso $Y = \{Y_n\}$ $\{\mathcal{F}_n\}$-predecible y uniformemente acotado. Para normalizar podemos suponer que la sucesión está acotada por **uno**.

Con ésta notación diremos que la transformada de una martingala por una sucesión multiplicadora es también una martingala.

§5. La desigualdad de Burkholder.

En lo que sigue vamos a tratar de dar un esbozo de cómo la desigualdad de Khintchine evolucionó en el contexto de la Teoría de Martingalas, para las demosraciones de éstos resultados remitimos a [19].

El siguiente resultado generaliza la desigualdad de Khintchine para variables aleatorias con una distribución arbitraria, para $p > 1$, (véase Marcinkiewicz-Zygmund [13]):

Teorema (Desigualdad de Marcinkiewicz-Zygmund): Sea $\zeta_1, \zeta_2, \ldots$, una sucesión de variables aletorias integrables e independientes con $\mathbf{E}[\zeta_i] = 0$, entonces para $p \geq 1$ existen contantes universales A_p y B_p (que dependen sólo de p), tal que

$$A_p \left\| \left(\sum_{j=1}^{n} \xi_j^2 \right)^{1/2} \right\|_p \leq \left\| \sum_{j=1}^{n} \zeta_j \right\|_p \leq B_p \left\| \left(\sum_{j=1}^{n} \xi_j^2 \right)^{1/2} \right\|_p , \qquad (20)$$

para todo $n \geq 1$.

Obsérvese que en los ejemplos 1) y 3) la sucesiones $\{X_n\}$ consideradas,

$$X_n = \sum_{k=1}^{n} \zeta_k,\ X_n = \sum_{k=1}^{n} a_k r_k \quad W_{2^n} = \sum_{k=1}^{2^n-1} c_k \psi_k \text{ y } H_{2^n} = c_0 + \sum_{n=1}^{2^n} \sum_{k=1}^{2^n-1} c_k^{(n)} \chi_k^{(n)} \text{ asi como } \sum_{j=1}^{n} \zeta_j$$

en el teorema anterior son martingalas. Una pregunta natural es entonces si la desigualdad de Khintchine, la desigualdad de Paley y la desigualdad de Marcinkiewicz-Zygmund se pueden extender a martingalas arbitrarias, la respuesta es positiva y es la famosa desigualdad de Burkholder (véase [2] y [4]).

Para establecer dicha desigualdad usaremos, tal como lo hicimos en la desigualdad de Paley y la desigualdad de Marcinkiewicz - Zygmund, la **función cuadrática de orden n y función cuadrática de X** , $S_n(X)$ y $S(X)$, que definimos como

$$S_n(X) = \left(\sum_{k=1}^{n} (\Delta X)_k^2 \right)^{1/2} \quad \text{y} \quad S(X) = \left(\sum_{k=1}^{\infty} (\Delta X)_k^2 \right)^{1/2},$$

respectivamente.

Teorema (Desigualdad de Burkholder): Sea $1< p < \infty$. Existen constantes positivas c_p y C_p, tal que si $X = \{X_n\}$ es una $\{\mathscr{F}_n\}$ - martingala, entonces para todo n

$$c_p \|S_n(X)\|_p \leq \|X_n\|_p \leq C_p\|S_n(X)\|_p. \tag{21}$$

ó en forma equivalente,

$$c_p \|S(X)\|_p \leq \|X\|_p \leq C_p\|S(X)\|_p \tag{21'}$$

La escogencia optimal de c_p satisface c\s(-1;p) = $O(p^{1/2} q)$ y $C_p = O(q^{1/2} p)$ donde $\frac{1}{p}+ \frac{1}{q} = 1$(de hecho, se tiene que $c_p = [18\, p^{3/2}/(p-1)]^{-1}$ y $C_p = 18\, p^{3/2}/(p-1)^{1/2}$).

La demostración original de Burkholder [2] utiliza de manera explícita el Teorema de Interpolación de Marcinkiewicz (véase Zygmund [21]). Más aún, como menciona D. Burkolder, Richard Gundy observó que la demostración de Paley de su desigualdad se puede adaptar al caso general de martingalas con modificaciones menores.

Mediante la desigualdad de Burkholder se puede obtener, por ejemplo, una Ley fuerte de los grandes números para martingalas y una extensión de la Ecuación de Wald.

Ahora bien, la desigualdad de Burkholder, se cumple para $p > 1$, mientras que la desigualdad de Marcinkiewicz - Zygmund se cumple también para el caso $p = 1$. ¿ Que pasa para el caso $p = 1$ en general? Como lo demuestra el siguiente ejemplo es falso.

Ejemplo: Sea $\zeta_1, \zeta_2, \ldots$ variables aleatorias independientes con distribución Bernoulli ($\mathbf{P}\{\zeta_1 = 1\} = \mathbf{P}\{\zeta_1 = -1\} = 1/2$) y sea

$$X_n = \sum_{j=1}^{n\wedge\tau} \zeta_j \quad \text{donde} \quad \tau = \inf\{n \geq 1: \sum_{j=1}^{n} \zeta_j = 1\}.$$

Luego $X = \{X_n\}$ es una es una martingala tal que

$$\|X_n\|_1 = \mathbf{E}[|X_n|] = 2\mathbf{E}[X_n^+] \to 2 \text{ si } n \to \infty \text{ y } \mathbf{E}[X_n] = 0,$$

es decir, $\{X_n\}$ es L^1 acotada pero

$$\|S_n(X)\|_1 = \mathbf{E}[S_n(X)] = \mathbf{E}\left(\sum_{j=1}^{n\wedge\tau} 1\right)^{1/2} = \mathbf{E}[\sqrt{n\wedge\tau}\,] \to \infty, \text{ si } n \to \infty. \blacksquare$$

Luego la desigualdad de Burkholder no es cierta para $p = 1$. Sin embargo, por la desigualdad maximal de Doob [5] (Corolario 3-1, Capítulo 3): para $p > 1$

$$A_p \|X_n\|_p \le \|X_n^*\|_p \le B_p \|X_n\|_p,$$

ella es equivalente a la siguiente desigualdad: para $1 < p < \infty$ existen constantes c_p y C_p tal que

$$c_p \|S_n(X)\|_p \le \|X_n^*\|_p \le C_p \|S_n(X)\|_p \qquad (22)$$

o en forma equivalente, para todo n

$$c_p \|S(X)\|_p \le \|X^*\|_p \le C_p \|S(X)\|_p\,, \qquad (22')$$

Es claro que para $p = 1$ (16) y (17) no son equivalentes, ya que en ese caso esa desigualdad de Doob no es cierta. La pregunta entonces es si para $p = 1$ (16) es todavía cierta. La respuesta es afirmativa y viene dada por el siguiente resultado :

Teorema (Davis): Sea $X = \{X_n\}_n$ una $\{\mathscr{F}_n\}$ - martingala, entonces existen constantes positivas c, C tal que

$$c\,\|S(X)\|_1 \le \|X^*\|_1 \le C\,\|S(X)\|_1 \qquad (23)$$

Burkholder [2] no sólo generaliza la primera desigualdad de Paley sino también la segunda desigualdad de Paley:

Teorema (Burkholder): Sea $X = \{X_n\}_n$ una $\{\mathscr{F}_n\}$ - martingala y sea $Y = \{Y_n\}$ una sucesión multiplicadora. Entonces dado $1 < p < \infty$ existe una constante c_p (independiente de X e Y) tal que para todo n

$$\|(Y\cdot X)_n\|_p \le c_p \|X_n\|_p \qquad (24)$$

es decir, en forma equivalente

$$\|(Y\cdot X)\|_p \leq c_p \|X\|_p \qquad (24')$$

Se puede probar que la mejor constante para ésta desigualdad es $p^* - 1$ donde $p^* = \max(p,q)$, con $\frac{1}{p} + \frac{1}{q} = 1$.

Proposición : La desigualdades de Burkholder (21) y (24) son equivalentes.

Demostración:

Si sabemos que $\|(Y\cdot X)_n\|_p \leq c_p \|X_n\|_p$ para todo n y cualquier sucesión multiplicadora Y, entonces para $t \in [0,1)$ fijo consideremos la martingala transformada $(Y \circ X)$ correspondiente a la sucesión multiplicadora $Y = \{r_n(t)\}_n$ con $\{r_n\}$ las funciones de Rademacher y $t \in [0,1)$, entonces por (24)

$$\|(Y\cdot X)_n\|_p^p = \|\sum_{k=1}^n r_k(t)\,\Delta X_k\|_p^p \leq C_p^p \|\sum_{k=1}^n \Delta X_k\|_p^p = C_p^p \|X_n\|_p^p$$

pero como $r_k^2 \equiv 1$, entonces X es una transformada de $(Y\cdot X)$ por la <u>misma</u> sucesión multiplicadora Y y así pués, de nuevo por (24)

$$\|X_n\|_p^p \leq C_p^p \|\sum_{k=1}^n r_k(t)\,\Delta X_k\|_p^p = C_p^p \|(Y\cdot X)_n\|_p^p ,$$

así pues

$$C_p^{-p} \|(Y\cdot X)_n\|_p^p \leq \|X_n\|_p^p \leq C_p^p \|(Y\cdot X)_n\|_p^p .$$

Ahora bien, integrando en $t\in[0,1]$ y aplicando el Teorema de Fubini, tenemos

$$\int_0^1 \|\sum_{k=1}^n r_k(t)\,\Delta X_k\|_p^p \, dt = \mathbf{E}\left[\int_0^1 |\sum_{k=1}^n r_k(t)\,\Delta X_k|^p \, dt\right]$$

y usando entonces la desigualdad de Khinchine se tiene

$$A_p^p \|S(X)\|_p^p = A_p^p \,\mathbf{E}\left[\sum_{k=1}^n (\Delta X_k)^2\right]^{p/2} \leq \mathbf{E}\left[\int_0^1 |\sum_{k=1}^n r_k(t)\,\Delta X_k|^p \, dt\right] \leq B_p^p \|S(X)\|_p^p$$

y por tanto obtenemos (21) :

$$C_p^{-1} A_p \|S(X)\|_p \leq \|X_n\|_p \; C_p B_p \|S(X)\|_p^p .$$

Para la implicación recíproca obsérvese que como la sucesión multiplicadora es uniformemente acotada por 1, entonces

$$S(Y\cdot X) = \Big[\sum_{k=1}^{n} (Y_k \Delta X_k)^2\Big]^{1/2} \le \Big[\sum_{k=1}^{n} (\Delta X_k)^2\Big]^{1/2} = S(X)$$

luego por (21), tenemos entonces (24)

$$\|(Y\cdot X)_n\|_p \le C_p \|S(Y\cdot X)\|_p = C_p \|S(X)\|_p \le C_p^{-1} C_p \|X\|_p . \blacksquare$$

§5. Generalizaciones de la desigualdad de Burkholder. Desigualdades del Buen λ.

Vamos a discutir ahora las posibles generalizaciones de la desigualdad de Burkholder :

En primer lugar, recuérdese que la desigualdad de Khintchine, que se puede considerar la fuente de inspiración de éstas desigualdades, es cierta para todo $p > 0$. Sin embargo , en todas las otras desigualdades el rango $0 < p < 1$ está excluido. La pregunta es ¿ que pasa en ese rango?, ¿ que tipo de restricciones son necesarias para obtener las desigualdades en ese rango?. El siguiente ejemplo, debido a Marcinkiewicz y Zygmund [20], demuestra que sin condiciones adicionales sobre la martingala, la desigualdad de Burkholder es falsa para el rango $0 < p < 1$.

Ejemplo:

Sea j un entero positivo y $\zeta_j = (\zeta_{j1}, \zeta_{j2},....)$, una sucesión de variables aleatorias independientes, con distribución

$$\mathbf{P}\{\zeta_{jk} = 1\} = 1 - (j+1)^{-1}, \quad \mathbf{P}\{\zeta_{jk} = -j\} = (j+1)^{-1}.$$

Sea $Z_j = (Z_{j1}, Z_{j2}, ...)$ una martingala definida por

$$Z_{jn} = \sum_{k=1}^{n} \zeta_{jk}.$$

En efecto Z_j es una martingala, ya que es la suma de variables aleatorias independientes centradas (véase el ejemplo 1).

Ahora veamos que poniendo $Z_{jn}^m = \sum_{k=1}^{n} \chi_{\{m>k\}} \zeta_{jk}$.

$$\lim_{m\to\infty} \lim_{j\to\infty} \|Z_j^m\|_p / \|S(Z_j^m)\|_p = \infty.$$

Si $j > 2n$, $\mathbf{P}\{\zeta_{jk} = -j\} < \mathbf{P}\{\zeta_{jk} = 1\}$ en consecuencia

$$\|Z_j^m\|_p = \sup_{1\leq n<\infty} \|Z_{jn}^m\|_p = \sup_{1\leq n<\infty} \| \sum_{k=1}^{n} \chi_{\{m>k\}} \zeta_{jk}\|_p = \sup_{1\leq n<\infty} \| \sum_{k=1}^{n\wedge m} \zeta_{jk}\|_p$$

$$= \sup_{1\leq n<\infty} \left(\int_\Omega | \sum_{k=1}^{n\wedge m} \zeta_{jk}|^p \, d\mathbf{P} \right)^{1/p} \geq \sup_{1\leq n<\infty} \left(\int_\Omega \left(\sum_{k=1}^{n\wedge m} \zeta_{jk} \right)^p d\mathbf{P} \right)^{1/p} = \sup_{1\leq n<\infty} (n\wedge m) = m.$$

De forma análoga se tiene:

$$\|S(Z_j^m)\|_p = \| (\sum_{k=1}^{m} (\zeta_{jk})^2)^{1/2}\|_p = \left(\int_\Omega \left| \sum_{k=1}^{m} (\zeta_{jk})^2 \right|^{p/2} d\mathbf{P} \right)^{1/p}$$

$$= \left(\int_{\{\zeta_{jk}=1\forall k\}} \left| \sum_{k=1}^{m} (\zeta_{jk})^2 \right|^{p/2} d\mathbf{P} + \int_{\{\delta_{jk}=1\forall k\}^c} \left| \sum_{k=1}^{m} (\zeta_{jk})^2 \right|^{p/2} d\mathbf{P} \right)^{1/p}$$

$$\leq \left[m^{p/2} (1 - \tfrac{1}{k+1})^m + (m\, j^2)^{p/2} (1 - (1 - \tfrac{1}{j+1})^m \right]^{1/p}$$

$$= \left[m^{p/2} ((1 - \tfrac{1}{j+1})^m + j^p (1 - (1 - \tfrac{1}{j+1})^m) \right]^{1/p}$$

como bien puede observarse:

$$\lim_{m\to\infty} \lim_{j\to\infty} \|Z_j^m\|_p = \infty \quad \text{y} \quad \lim_{m\to\infty} \lim_{j\to\infty} \|S(Z_j^m)\|_p = \infty$$

con $0 < p < 1$, de tal manera que no podemos tener una desigualdad del tipo

$$\|Z\|_p \leq C\|S(Z)\|_p.$$

En segundo lugar, queremos generalizaciones de la desigualdad de Burkholder para funciones más generales que $\Phi(\lambda) = \lambda^p$, así como también para una clase general de operadores sobre martingalas (que por supuesto incluya a los operadores $(\ )^*$ y S). D. Burkholder y R. Gundy trabajaron en éstas generalizaciones en su famoso artículo *Extrapolation and interpolation of quasilinear operators on martingales* publicado en Acta Mathematica [3] (véase tambíen [4]). Vamos a dsicutir en detalle aca la generalizaciòn a de la desigualdad de Burkholder para funciones Φ más generales.

Es claro que si tuviéramos desigualdadades entre funciones de distribución, entonces mediante el argumento de interpolación de Marcinkiewicz se podrían obtener facilmente la desigualdades integrales. Sin embargo tales desigualdades entre funciones de distribución no siempre ocurren para todo $\lambda > 0$ para los operadores que nos interesan, como muestra el siguiente ejemplo.

Ejemplo: Sea $X_n = \sum_{k=1}^{n} \frac{r_k(t)}{k}$ donde $\{r_k\}$ son la funciones de Rademacher, tenemos entonces $S(X) = \left(\sum_{k=1}^{\infty} \frac{1}{k^2}\right)^{1/2}$. Por la desigualdad de Khintchine sabemos que

$$A_p \left(\sum_{k=1}^{\infty} \frac{1}{k^2}\right)^{1/2} \leq \left\| \sum_{k=1}^{n} \frac{r_k}{k} \right\|_p \leq B_p \left(\sum_{k=1}^{n} \frac{1}{k^2}\right)^{1/2},$$

pero $m\{ t\in[0,1): S(X) > \lambda\} = 0$, para todo λ suficientemente grande, del otro lado para n un entero cualquiera dado

$$m\{X^* > \lambda\} \geq m\left\{t\in[0,1): \left| \sum_{k=1}^{n} \frac{r_k(t)}{k} \right| > \lambda\right\} > 0,$$

para todo $\lambda > 0$ por lo que una desigualdad entre sus funciones de distribución no es posible.

A pesar de éste hecho, Burkholder y Gundy en [3] caracterizan aquellos λ para los cuales las desigualdades entre funciones de distribución se cumplen, lo que les permite, luego de justificar que tales λ son suficientes para integrar, obtener fórmulas integrales. Esta es la **"condición del buen λ"**: $\lambda > 0$ es un **buen λ** para una variable aletoria Y si existen $\alpha > 0$ y $\beta > 1$ tal que

$$\mathbf{P}\{Y > \lambda\} \leq \alpha\, \mathbf{P}\{Y > \beta\, \lambda\}.$$

Sin embargo en [4], Burkholder establece que usando desigualdades más débiles que las anteriores (pero válidas para todo λ) y que hoy es lo que se conocen como **desigualdades del buen λ**, se pueden obtener también las desigualdades integrales deseadas (véase la próxima Proposición). Así pues, las desigualdades del buen λ son las traducción correcta, al campo de las Probabilidades, del argumento de Interpolación de Marcinkiewicz.

El próximo resultado explica detalladamente este argumento dado por Burkholder en [4]. Pero primero especifiquemos el tipo de funciones Φ para las cuales queremos generalizar la desigualdad de Burkholder.

Definición : Una función Φ no idénticamente nula, no-decreciente, continua sobre $[0,\infty]$ con $\Phi(0) = 0$ y que satisface la condición de crecimiento

$$\Phi(2\lambda) \leq C\,\Phi(\lambda) \tag{25}$$

se dice que Φ es una **función de Young**.

Para eliminar el caso trivial, supongamos que Φ no es idénticamente nula.

Es claro que para $0 < p < \infty$ $\Phi(\lambda) = \lambda^p$ es una función de Young.

Como propiedades adicionales que seran, eventualmente, necesarias tenemos la siguientes:

i) $\Phi(\lambda_1 \vee \lambda_2) \leq \Phi(\lambda_1) + \Phi(\lambda_2)$

ii) $\Phi(\lambda_1 + \lambda_2) \leq \Phi(2\lambda_1) + \Phi(2\lambda_2) \leq c\,[\Phi(\lambda_1) + \Phi(\lambda_2)\,]$

Proposición : Supongamos que X e Y son variables aleatorias no-negativas y sobre un espacio de probabilidad $(\Omega, \mathscr{F}, \mathbf{P})$ y $\beta > 1$, $\delta > 0$, $\varepsilon > 0$, son números reales, tales que, para todo $\lambda > 0$

$$\mathbf{P}\{Y > \beta\lambda,\ X \leq \delta\lambda\} \leq \varepsilon\ \mathbf{P}\{Y > \lambda\} \tag{26}$$

Sea Φ una función de Young y sean $\gamma > 0$ y $\eta > 0$, números reales tales que para cualquier $\lambda > 0$,

$$\Phi(\beta\lambda) \leq \gamma\,\Phi(\lambda)\ ,\ \ \Phi(\delta^{-1}\lambda) \leq \eta\,\Phi(\lambda), \tag{27}$$

Finalmente, supongamos que $\gamma\,\varepsilon < 1$. Entonces

$$\mathbf{E}[\Phi(Y)] \leq \frac{\gamma\eta}{1-\gamma\varepsilon}\ \mathbf{E}\,[\Phi(X)]. \tag{28}$$

Observación: La existencia de γ y η cumpliendo con (27) es garantizada por (25). Por ejemplo, una posible escogencia para γ es C^{k_0} donde k_0 es un entero positivo, tal que $2^{k_0-1} < \beta \leq 2^{k_0}$ y C es la constante en (25)

$$\Phi(2^{k_0-1}\lambda) < \Phi(\beta\lambda) \leq \Phi(2^{k_0}\lambda) \leq c\,\Phi(2^{k_0-1}\lambda) \leq \ldots \leq C^{k_0}\Phi(\lambda),$$

y análogamente para el caso de η.

Demostración:

Nosotros podemos asumir en la prueba que **E** $[\Phi(Y)]$ es finito, pues si Y satisface (26) entonces $Y\wedge n$ lo satisface también:

$$\mathbf{P}\{Y\wedge n > \beta\lambda, X \leq \delta\lambda\} = \mathbf{P}\{Y\wedge n > \beta\lambda, Y \leq n, X \leq \delta\lambda\} + \mathbf{P}\{Y\wedge n > \beta\lambda, Y > n\ X \leq \delta\lambda\}$$

$$\leq \mathbf{P}\{Y > \beta\lambda, Y \leq n, X \leq \delta\lambda\} + \mathbf{P}\{n > \beta\lambda, Y > n, X \leq \delta\lambda\}$$

$$\leq 2\mathbf{P}\{Y > \beta\lambda, X \leq \delta\lambda\} \leq 2\varepsilon\,\mathbf{P}\{Y > \lambda\} \leq 2\varepsilon\,\mathbf{P}\{Y\wedge n > \lambda\}.$$

A su vez si (28) ocurre para $Y\wedge n$, $n \geq 1$, entonces ocurrirá para Y., ya que, por el Teorema de convergencia monótona se tiene:

$$\int_\Omega \Phi(Y)\, d\mathbf{P} = \lim_{n\to\infty} \int_\Omega \Phi(Y\wedge n)\, d\mathbf{P} \leq \frac{\gamma\eta}{1-\gamma\varepsilon} \int_\Omega \Phi(X)\, d\mathbf{P}.$$

Consideremos ahora la medida de Lebesgue - Stieltjes para Φ, denotada por μ_Φ, definida como

$$\mu_\Phi([a,b)) = \int_a^b d\Phi(\lambda) = \Phi(b) - \Phi(a),\ \ 0 \leq a < b \leq \infty.$$

Esta medida es positiva y σ-finita sobre $\mathcal{B}[0,\infty)$. Si Z es una variable aletoria no-negativa sobre $(\Omega, \mathcal{F}, \mathbf{P})$, entonces

$$\Phi(Z) = \int_0^Z d\Phi(\lambda) = \int_0^\infty \chi_{\{Z>\lambda\}}\, d\Phi(\lambda).$$

Usando el teorema de Fubini, obtenemos entonces

$$\mathbf{E}[\Phi(Z)] = \int_\Omega \left(\int_0^\infty \chi_{\{Z>\lambda\}}(\lambda)\, d\Phi(\lambda)\right) d\mathbf{P} = \int_0^\infty \left(\int_\Omega \chi_{\{Z>\lambda\}}(\omega)\, d\mathbf{P} \right) d\Phi(\lambda) = \int_0^\infty \mathbf{P}\{Z>\lambda\}$$

$$d\Phi(\lambda)$$

La condición (26) implica que, para todo $\lambda > 0$:

$$\mathbf{P}\{Y > \beta\lambda\} = \mathbf{P}\{Y > \beta\lambda, X \leq \delta\lambda\} + \mathbf{P}\{Y > \beta\lambda, X > \delta\lambda\} \leq \varepsilon\,\mathbf{P}\{Y > \lambda\} + \mathbf{P}\{X > \delta\lambda\},$$

Entonces, tenemos

$$\mathbf{E}\,[\Phi(\beta^{-1}\,Y)] = \int_0^\infty \mathbf{P}\{\beta^{-1}Y > \lambda\}\,d\Phi(\lambda) = \int_0^\infty \mathbf{P}\{Y > \beta\lambda\}\,d\Phi(\lambda)$$

$$\leq \int_0^\infty \varepsilon\,\mathbf{P}\{Y > \lambda\}\,d\Phi(\lambda) + \int_0^\infty \mathbf{P}\{X > \delta\lambda\}\,d\Phi(\lambda) = \varepsilon\mathbf{E}[\Phi(Y)\,] + \mathbf{E}[\Phi(\delta^{-1}\,X)]$$

$$\leq \varepsilon\,\mathbf{E}\,[\Phi(Y)\,] + \eta\,\mathbf{E}\,[\Phi(X)\,],$$

como

$$\mathbf{E}\,[\Phi(Y)\,] = \mathbf{E}\,[\Phi(\beta\beta^{-1}\,Y)] \leq \gamma\,\mathbf{E}\,[\Phi(\beta^{-1}\,Y)] \leq \gamma\,\varepsilon\,\mathbf{E}\,[\Phi(Y)\,] + \gamma\,\eta\,\mathbf{E}\,[\Phi(X)\,],$$

$$(1 - \gamma\,\varepsilon)\,\mathbf{E}\,[\Phi(Y)\,] \leq \gamma\,\eta\,\mathbf{E}\,[\Phi(X)],$$

y dado que $\mathbf{E}\,[\Phi(Y)\,] < \infty$, se tiene

$$\mathbf{E}\,[\Phi(Y)\,] \leq \frac{\gamma\eta}{1 - \gamma\,\varepsilon}\,\mathbf{E}\,[\Phi(X)]\;. \blacksquare$$

Observaciones

i) Si $\mathbf{P}\{Y < \lambda\} \leq C\,\mathbf{P}\{X > \delta\lambda\}$, entonces (28) es inmediato, para cualquier Φ una función de Young

$$\mathbf{E}\,[\Phi(Y)\,] = \int_\Omega \Phi(Y)\,d\mathbf{P} = \int_0^\infty \mathbf{P}\{Y > \lambda\}\,d\Phi(\lambda) \leq C \int_0^\infty \mathbf{P}\{X/_\delta > \lambda\}\;d\Phi\,(\lambda)$$

$$= C\,\mathbf{E}\,[\Phi(\tfrac{X}{\delta})] \leq C'\,\mathbf{E}\,[\Phi(X)\,].$$

ii) Si λ es un " buen λ" para Y, entonces ésta desigualdad y (27) implican

$$\mathbf{P}\{Y > \beta\lambda\} \leq \mathbf{P}\{Y > \beta\lambda,\;\; X \leq \delta\lambda + \mathbf{P}\{Y > \beta\lambda,\;\; X > \delta\lambda\} \leq \varepsilon\,\mathbf{P}\{Y > \lambda\} + \mathbf{P}\{X > \delta\lambda\}$$

$$\leq \alpha\varepsilon\,\mathbf{P}\{Y > \beta\lambda\} + \mathbf{P}\{X > \delta\lambda\},$$

por tanto, si $\alpha\varepsilon < 1$, tenemos

$$\mathbf{P}\{Y > \beta\lambda\} \leq C\,\mathbf{P}\{X > \delta\lambda\},$$

con $C = \frac{1}{1-\alpha\varepsilon}$ y volvemos al caso anterior.

El problema en el ejemplo de Marcinkiewicz y Zygmund se debe a que la martingala puede tener saltos muy grandes. La idea crucial de las desigualdades del buen λ es que cuando una martigala cruza el nivel λ, no puede irse demasiado lejos, es decir la excursión está acotada superiormente por un múltiplo de λ. Esta idea aparece ya en artículos de R. Gundy de fines de los años 60 (véase [7]) .

Tenemos ahora que si queremos generalizar la desigualdad de Burkholder para que sea cierta para toda función de Young, sabemos, por el ejemplo de de Marcinkiewicz y Zygmund, que para que ello sea posible hay que suponer condiciones adicionales sobre la martingala, es decir restringir el conjunto de martingalas que se está considerando, en ese sentido vamos a considerar la siguiente clase de martingalas:

Definición : Sea $X = \{X_n\}$ una $\{\mathcal{F}_n\}$ - martingala. X se dice **regular** si

$$\text{i) } \mathbf{E}\left[\,|\Delta X_k|\,|\mathcal{F}_{k-1}\right] \geq \delta, \tag{29}$$

$$\text{ii) } \mathbf{E}\left[\,(\Delta X_k)^2\,|\mathcal{F}_{k-1}\right] = 1, \tag{30}$$

para todo $k \geq 1$.

En [7] R. Gundy esteblece ya ciertas condiciones de regularidad sobre martingalas (la Condición (MZ)) para tratar problemas de convergencia c.s.

Finalmente estamos listos para enunciar la generalización de la desigualdad de Burkholder (22) para funciones de Young:

Teorema (Burkholder - Gundy): Sea Φ una funcion de Young, $X = \{X_n\}$ una $\{\mathcal{F}_n\}$ - martingala regular, $Y = \{Y_n\}$ una sucesión multiplicadora. Entonces existen constantes universales c y C tal que

$$c\,\mathbf{E}\,[\Phi(S(Y\circ X))\,] \leq \mathbf{E}[\Phi((Y\circ X)^*)] \leq C\,\mathbf{E}\,[\Phi(S(Y\circ X))\,]. \tag{31}$$

La escogencia de c y C, depende sólo de Φ y de la regularidad de X.

Por lo expuesto anteriormente para obtener dichas desigualdades es suficiente probar las siguientes desigualdades del buen λ para los operadores S y ( )* :

Proposición : Sea $X = \{X_n\}$ una $\{\mathcal{F}_n\}$ - martingala regular, $Y = \{Y_n\}$ una sucesión multiplicadora. Entonces si $\beta > 1$ y $0 < \delta < \beta - 1$, se tiene, para todo $\lambda > 0$

$$\mathbf{P}\{(Y\cdot X)^* > \beta\lambda,\ S(Y\cdot X) \vee Y^* \leq \delta\lambda\} \leq \frac{c\delta^2}{(\beta-\delta-1)^2}\mathbf{P}\{(Y\cdot X)^* > \lambda\}, \qquad (32)$$

y

$$\mathbf{P}\{S(Y\cdot X) > \beta\lambda,\ (Y\cdot X)^* \vee Y^* \leq \delta\lambda\} \leq \frac{c\delta^2}{(\beta-\delta-1)}\mathbf{P}\{S(Y\cdot X) > \lambda\}. \qquad (33)$$

Otra manera de controlar los saltos de las martingalas para obtener desigualdades integrales, sin pedir la regularidad, es la siguiente

Teorema: Sea Φ una función de Young $X = \{X_n\}$ una $\{\mathcal{F}_n\}$ - martingala, tal que

$$|\Delta X_n| \leq W_n \in \mathcal{F}_{n-1}, \qquad (34)$$

es decir, X tiene incrementos $\mathcal{F}_{n-1}$ - acotados. Entonces existe una constante c, que depende sólo de Φ, tal que

i) $\mathbf{E}[\Phi(X^*)] \leq c\,\mathbf{E}[\Phi(S(X))] + c\,\mathbf{E}[\Phi(W^*)]$ (35)

ii) $\mathbf{E}[\Phi(S(X))] \leq c\,\mathbf{E}[\Phi(X^*)] + c\,\mathbf{E}[\Phi(W^*)]$. (36)

De nuevo por la proposición vista basta probar un par de desigualdades del buen λ:

Proposición : Sea $X = \{X_n\}$ una $\{\mathcal{F}_n\}$-martingala, tal que $|\Delta X_n| \leq W_n \in \mathcal{F}_{n-1}$, entonces, si $\beta > 1$ y $0 < \delta < \beta - 1$, se tiene, para todo $\lambda > 0$

$$\mathbf{P}\{X^* > \beta\lambda,\ S(X) \vee W^* \leq \delta\lambda\} \leq \frac{2\delta^2}{(\beta-\delta-1)^2}\ \mathbf{P}\{X^* > \lambda\} \qquad (37)$$

y

$$\mathbf{P}\{S(X) > \beta\lambda,\ X^* \vee W^* \leq \delta\lambda\} \leq \frac{2\delta^2}{(\beta^2-\delta^2-1)}\ \mathbf{P}\{S(X) > \lambda\} \qquad (38)$$

Por otra parte, si pedimos no negatividad para la martingala, tenemos todavía el siguiente resultado:

Teorema: Sea Φ una funcion de Young y $X = \{X_n\}$ una $\{\mathcal{F}_n\}$ - martingala no negativa. Entonces, existe una constante C, que depende sólo de Φ, tal que

$$\mathbf{E}\,[\Phi(S(X))\,] \le C\ \mathbf{E}[\Phi(X^*)] \tag{39}$$

Como ya es usual para probar éste Teorema, de nuevo por la Proposición vista, basta obtener la siguiente desigualdad del buen λ,

Proposición : Sea $X = \{X_n\}$ una $\{\mathcal{F}_n\}$ - martingala no negativa. Entonces si $\beta > 1$ y $0 < \delta < \beta - 1$, se tiene para todo $\lambda > 0$

$$\mathbf{P}\{S(X) > \beta\lambda\,,\ X^* \le \delta\lambda\} \le \frac{2\delta^2}{(\beta^2-\delta^2-1)}\ \mathbf{P}\{S(X) > \lambda\} \tag{40}$$

La otra posible manera de generalizar la desigualdad de Burkholder es no restringir el conjunto de las martingalas sino el conjunto de las funciones Φ a considerar. Por ejemplo vamos a probar que en el caso que Φ es convexa tenemos la generalización (obsérvese que si $p > 1$, la función $\Phi(\lambda) = \lambda^p$ es convexa).

Primero necesitamos discutir algunos resultados sobre funciones convexas. Para simplificar suponemos que la función convexa Φ se puede escribir como

$$\Phi\,(\lambda) = \int_0^\lambda \phi(\gamma)\, d\gamma, \tag{41}$$

con ϕ una función <u>estrictamente</u> creciente y no negativa en $[0,\infty)$, (en general, ésta representación es cierta pero sólo podemos decir que ϕ es monótona creciente , véase Wheeden y Zygmund [19], Capítulo 7 §6) .

A Φ se le puede asociar una función convexa Ψ del mismo tipo (su conjugada en el sentido de Young) tal que

$$\Psi\,(\lambda) = \int_0^\lambda \psi(\gamma)\, d\gamma, \tag{41'}$$

donde $\psi = \phi^{-1}$, la función inversa de ϕ. Es fácil de ver que

i) $\lambda\,\phi(\lambda) = \Phi(\lambda) + \Psi(\phi(\lambda))$,

ii) $\int_0^\gamma \lambda d\psi(\lambda) = \Phi(\psi(\gamma))$.

iii) Desigualdad de Young: $\lambda\gamma \leq \Phi(\lambda) + \Psi(\gamma)$

iv) $\Phi(\frac{\lambda}{\alpha}) \leq \frac{1}{\alpha}\,\Phi(\lambda)$, para todo $\alpha \geq 1$.

Más aún, de (20), tomando $\rho = \sup_{\lambda>0} \left(\frac{\lambda\phi(\lambda)}{\Phi(\lambda)} \right)$, se obtiene fácilmente $1 < \rho \leq C - 1 < \infty$.

v) para todo $\alpha > 1$

$$\Phi(\alpha\lambda) \leq \alpha^{\rho}\Phi(\lambda)$$

y

$$\Psi(\lambda) \leq (\rho - 1)\,\Phi(\psi(\lambda)).$$

Por otra parte, utilizando (15'), Teorema 3' del Capítulo 3, para ψ y usando ii) obtenemos

$$\mathbf{E}\,[\,\Phi(\psi(X_n^*))\,] \leq \mathbf{E}\,[\,X_n\,\psi(X_n^*)\,].$$

Luego, por la desigualdad de Young,

$$\mathbf{E}\,[\,\Phi(\psi(X_n^*))\,] \leq \mathbf{E}\,[\,\Phi(\psi(\frac{X_n^*}{\rho}))\,] + \mathbf{E}\,[\,\Psi(\rho\, X_n)\,],$$

y por iv)

$$\frac{\rho-1}{\rho}\,\mathbf{E}\,[\,\Phi(\psi(X_n^*))\,] \leq \mathbf{E}\,[\,\Psi(\rho\, X_n)\,].$$

Usando, finalmente, v) obtenemos

$$\mathbf{E}\,[\,\Psi(X_n^*)\,] \leq \rho\,\mathbf{E}\,[\,\Psi(\rho\, X_n)\,]. \qquad (42)$$

Obsérvese que en el caso $\Psi(\lambda) = \lambda^p$, $p > 1$, (40) es precisamente la desigualdad de Doob [5].

Para establecer la generalización de la desigualdad de Burkholder en éste caso necesitamos se utiliza la siguiente desigualdad que tiene importancia en si misma

Proposición (Burkholder-Davis-Gundy)**:** Sea Φ una función de Young convexa. Sea $\{\zeta_n\}$ una suceción devariables aletorias no negativas en $(\Omega,\mathcal{F},\mathbf{P})$, entonces

$$\mathbf{E}\left[\Phi\left(\sum_{k=1}^{\infty} \mathbf{E}[\zeta_k|\mathcal{F}_{k-1}]\right)\right] \leq C\mathbf{E}\left[\Phi\left(\sum_{k=1}^{\infty} \zeta_k\right)\right] \tag{43}$$

Veamos entonces la generalización de la desigualdad de Burkholder para funciones convexas:

Teorema (Burkholder-Davis-Gundy): Sea Φ una funcion de Young convexa, sea $X = \{X_n\}$ una $\{\mathcal{F}_n\}$ - martingala . Entonces existen constantes universales c y C tal que

$$c\,\mathbf{E}[\Phi(S(X))] \leq \mathbf{E}[\Phi(X^*)] \leq C\,\mathbf{E}[\Phi(S(X))]. \tag{44}$$

La escogencia de c y C, depende sólo de Φ.

La generalización de la desigualdad de Burkholder - Davis - Gundy para una clase general de operadores definidos sobre la clase $\mathcal{C}$ de todas transformadas de martingalas regulares, como ya mencionamos, fue realizado por D. Burkholder y R. Gundy un su artículo de Acta Matemática [3]. Sin embargo, las demostraciones en ese artículo, si bien contienen todas las ideas básicas a las desigualdades del buen λ, no las utilizan con toda su potencia, por lo que las demostraciones son muy complicadas. Precisamente con la formulación de las desigualdades del buen λ, y la técnica que ellas implican dichas demostraciones se simplifican de manera notable.

Por otra parte, las desigualdades del buen λ son una técnica muy poderosa para obtener desigualdades integrales e incluso estimaciones locales ó puntuales de un operador, que lamentablemente es poco conocida en Probabilidades. Sin embargo ellas se ha se han utilizado en Análisis de manera muy efectiva para establecer, por ejemplo, continuidad de operadores de Calderón- Zygmund en espacios $L^p(\omega)$ para un peso de Muckenhoupt ω, ó la ya mencionada caracterización de los espacios H^p .

Finalmente, la desigualdad de Khintchine ha tenido una influncia descisiva en otras áreas de las Matemáticas, como es el caso en Análisis de la Teoría de Littlewood- Paley y de la reciente Teoría de Ondículas que por razones obvias no discutiremos en este ensayo.

BIBLIOGRAFIA

[1] Billingsley, P. *Probability and Measure.* Wiley, N.Y. 1986.

[2] Burkholder, D.L. *Martingale Transform.* Ann. Math. Stat. 37 (1966) 1494-1504.

[3] Burkholder, D.L., Gundy, R. *Extrapolation and interpolation of quasilinear operators on martingales.* Acta Math.124 (1970) 249-324.

[4] Burkholder, D.L. *Distribution function inequalities for martingales.* Ann.Prob.1 (1973) 19-42.

[5] Doob, J.L. *Stochastic Processes.* Wiley, N.Y. 1953.

[6] Gundy, R. *Martingale Theory and pointwise convergence of certain ortogonal series.* Trans. Am. Math. Soc. 124 (1966) 228- 248.

[7] Gundy, R. *The Martingale version of a Theorem of Marcinkiewicz and Zygmund.* Ann. Math. Statist., 38 (1967), 725-734.

[8] Gundy, R. *On a class of Martingale series.* Proc. of the Conference on Orthogonal expansions and their continuos analogues. Southern Ilinois Univ. Edwardsville (1967). 99 - 101.

[9] Hunt, R. *Almost everywhere convergence of Walsh-Fourier series of L^2 functions.* Actes. Inter. Congress Math. 2 (1970) 655 - 661.

[10] Kac, M. *Statistical independence in Probability, Analysis and Number Theory.* Carus Math Monogr. # 12 MAA

[11] Kaczmarz, S. & Steinhaus, H. *Le systeme orthogonal de M. Rademacher*. Studia Mathematica 2 (1930) 231-247.

[12] Marcinkiewicz, J. *Quelques theorémes sur les séries orthogonales.* Ann. Soc. Polon. Math. 16 (1937) 84-96.

[13] Marcinkiewicz, J., Zygmund, A. *Quelques théoremes sur les fonctions indépendantes.* Studia Math.7 (1938) 104 - 120.

[14] Paley, R.E.A.C. *A remarkable series of orthogonal functions I.* Proc. London Math. Soc. 34 (1932) 241-264.

[15] Rademacher, H. *EinigenSatze ueber Reihen von allgemeinenen Orthogonalfunktionen.* Math. Ann. vol 87 (1922) 112-138.

[16] Schipp,F., Wade, W.R. & Simon, P. *Walsh Series: an introduction to dyadic Harmonic Analysis.* Akadémiai Kiadó, Budpest (1990).

[17] Stein, E .*The development of square functions in the work of A. Zygmund.* Conference on Harmonic Analysis in honor of Antoni Zygmund,W. Beckner,A.P. Calderón, R. Fefferman and P.W. Jones (Eds.) Wadsworth,Belmont,Calif.(1982).

[18] Szareck, S.J. *On the best constants in the Khintchine inequality.* Studia Math. LVIII(1976) 197-208.

[19] Urbina,W. *Teoría de Martingalas y Aplicaciones.* V Escuela Venezolana de Matemáticas, Mérida 1992.

[20] Wheeden, R., Zygmund, A. *Measure and Integral. An Introduction to Real Analysis.* Marcel Dekker, Inc.New York 1977.

[21] Zygmund, A. *Trigonometric Series.* 2 ed. Cambridge 1959.

SOME PROBLEMS INVOLVING DILATION EGUATIONS.

DICK GUNDY

Departament of Mathematics

Rutgers University

Mérida, 27 de Enero de 1994.-

I do not know where the term dilation equation originated. The concept has become very well-known in the past few years in the research on "wavelets" in connection with the notion of a multi-resolution analysis.The idea behind the construction of a multi-resolution analysis[1] is the following:

Suppose we are given a sequence of subspaces $V_j \subset V_{j+1}$ of $L^2(\mathbb{R})$, $j : -\infty < j < \infty$, with $\bigcap_j V_j = \emptyset$ and $\bigcup_j V_j \subset_{densely} L^2(\mathbb{R})$. Suppose further that there exists a function $\varphi(x)$ such that the sums

$$\| \sum a_k \varphi(x-k) \|_2^2 \cong (\sum a_k^2)$$

where $\cong$ means a double inequality of the form:

$$A(\sum a_k^2) \leq \| \sum a_k \varphi_k \|_2^2 \leq B(\sum a_k^2)$$

where A, B do not depend on the sequence $\{a_k\}$ and

$$V_0 = \{\sum a_k \varphi(x-k); \sum a_k^2 < \infty\}.$$

In the language of the subject, V_0 has a "Riez basis" formed with the translates of $\varphi(x)$ by $k \in \mathbf{Z}$.

Further, suppose we define

$$\varphi_1(x) = 2\varphi(2x)$$

and

$$\varphi_n(x) = 2^n \varphi(2^n x), \quad n \in \mathbf{Z}.$$

[1]=m.r.a.

The function $\varphi_n(x)$ is a <u>dilation</u> of φ, the scale being 2^{-n}. (We should write

$$\varphi_t(x) = \frac{1}{t}\varphi(\frac{x}{t}),$$

but I prefer the previous notation for dyadic dilations, and I want the scale $2^{-n} \to 0$ as $n \to +\infty$ (rather than $-\infty$).)

Continuing in the specification of a m.r.a., we suppose that

$$V_1(\supset V_0) := \{\sum b_n \varphi_1(x - k/2)\}$$

so that the dilations (by 2^{-1}), and translations by half-integers, generate V_1. That is, every function in V_1 has an expression of the above form. In particular since $\varphi(x)(= \varphi_0(x))$ is in V_0, we may find $b_k, k = 0, \pm 1, \pm 2, \ldots$ such that

$$\varphi_0(x) = \sum b_k \varphi_1(x - k/2).$$

Such a φ_0 satisfies a dilation equation (or a two-scale resolution equation : the terminology varies, and I have adopted the term from G. Strang, Bull. Amer. Math. Soc. articles on wavelets. See 1990-93).

There are many real-life examples of m.r.a. One of these is the cardinal spline family.

Let

$$\chi(x)_{[0,1]} = \begin{cases} 1 & \text{if } 0 \le x \le 1 \\ 0 & \text{otherwise} \end{cases}$$

$$\varphi^{(2)}(x) = (\chi_{[0,1]} * \chi_{[0,1]})(x),$$

where $*$ means convolution.

$$\varphi^{(n)}(x) = (\chi_{[0,1]} * \cdots * \chi_{[0,1]})(x)$$

where we have indicated an n-fold convolution.

It can be shown (see "Wavelets": Chui) that for each fixed $n, \varphi^{(n)}(x)$ generates (by translations ($\mathbf{Z}$) and dilations ($2^n, n \in \mathbf{Z}$)) a m.r.a..

I left out a particularly interesting case: $\varphi^{(1)}(x) = \chi_{[0,1]}(x)$. The m.r.a. here is generated by the Haar functions, and the projections of $f \in L^2(\mathbb{R})$ on $V_j, j \in \mathbf{Z}$ form a martingale. It was this coincidence that attracted me to the problem of studying m.r.a., from the same perspective as I had done with martingales. The analogy is strong, but there are important differences. Let me recall for you what is known about the Haar functions from a martingale point of view:

Let E_j be the projection of a function $f \in L^2(\mathbb{R})$ onto the subspace V_j. Further, suppose that f is supported on the unit interval, so that the total measure of the support is one. That is, we consider a function $f \in L^2([0,1])$. The sequence $E_j, j = 0, 1, 2, \ldots$ has an interesting local property: the differences

$$E_j - E_{j-1}(f) = E_j(f) - E_{j-1}(f)$$

are mutually orthogonal in a strong sense: If we integrate the product of differences:

$$\int (E_j - E_{j-1})(f)(E_k - E_{k-1})(f)\chi_A dx = 0$$

if $j \neq k$ and A is any dyadic interval of length 2^l, where $l = \min(k, j)$. (Dyadic interval:

$$\{x : m - 1/2^l \leq x \leq m/2^l\}.)$$

This is the martingale property of the sequence $\{E_j(f)\}$ in this context. The generality of the set A is the key here: not only are the differences $(E_j - E_{j-1})(f)$ orthogonal, but they are even orthogonal when restricted to "small" sets A. This has an important consequence, as follows. Let

$$S^2(f)(x) = \sum((E_j - E_{j-1})(f))^2(x)$$

and

$$f^*(x) = \sup |E_j(f)|(x).$$

The two functionals $S(f)$ and f^* clearly reflect different properties of the sequence $E_j(f) : S(f)$ measures oscillations, while f^* reflects the maximum rise or fall of the sequence. For a general numerical sequence, there is clearly no relation between $\sup_n |a_n|$ and $\sum(a_n - a_{n-1})^2$. However for the sequence of projections $E_j(f)$, the martingale theory tells us that

$$\|S(f)\|_p \cong \|f^*\|_p$$

for all $0 < p < \infty$. But $\|f\|_\infty \cong \|S(f)\|_\infty$ is false. However, the L_p equivalences are based on a distribution function inequality that resembles

$$P(S(f) > \lambda) \cong P(f^* > \lambda).$$

This inequality is not true, since it would imply that the L_∞ norms are not comparable. However, the inequality does hold for "enough" lambdas. This fact allows us to make integrations of the kind

$$\int_0^\infty \alpha'(\lambda) P(S(f) > \lambda) d\lambda \cong \int_0^\infty \alpha'(\lambda) P(f^* > \lambda) d\lambda$$

where the $\alpha'(\lambda) \geq 0$, and

$$\alpha(\lambda) = \int_0^\lambda \alpha(\gamma) d\gamma$$

satisfies

$$\alpha(2\lambda) \leq c\alpha(\lambda).$$

Such distribution function comparisons are known as "good λ" inequalities. In addition to integral inequalities, they allow us to assert that

$$P(S(f) < \infty) = P(f^* < \infty)$$

(or more precisely,

$$\{S(f) < \infty\} =_{a.e.} \{f^* < \infty\}).$$

All this can be generalized for martingales, and is independent of dilation-translation considerations. This being so, it is of interest to ask whether there is a similar pair of functionals $S(f)$ and f^* which admit a comparison of the kind we have discussed, outside of the martingale contex, yet within the scope of dilations and translation approximation procedures, such as the multiresolution analysis discussed above.

This question, concerning m.r.a.s, we cannot answer. However, there is yet another setting, where translations and dilations come into play, where we can find similar phenomena. The first situation of this nature is the study of harmonic functions in $\mathbb{R}^n_+$ or $\mathbb{R}^1_+$, for simplicity.

Let

$$\mathbb{R}^1_+ = \{(x,t) : -\infty < x < \infty, t > 0\}$$

and

$$1/t\varphi(x/t) := \varphi_t(x).$$

Define

$$\Phi_t(f)(x) = \int_{-\infty}^{\infty} f(y)\varphi_t(x-y)dy$$

where

$$\varphi(x) = \frac{1}{\pi}\frac{1}{1+x^2}$$

With $\varphi(x)$ defined in this way, $\Phi_t(f)(x)$ is harmonic in $(x,t) : \Delta\Phi \equiv 0$ and $\Phi_t(f)(x)$ is the harmonic extension of f, from $\mathbb{R}^1$ to $\mathbb{R}^1_+$. For some time, we were under the spell of Laplace's equation, and did not consider that $\Phi_t(f)(x)$ was formed by a dilation kernel (which also satisfied a differential equation). However, the functions $\Phi_t(f)(x)$ behave, in many ways, like martingales, and in particular, like the Haar series projections E_j, as $t \to 0$. Here, the dyadic parameter 2^{-n} is replaced by the continuous dilation t, with the multiplicative Haar measure dt/t replacing the unit weights in the sum defining $S(f)$. (Note that

$$\Delta 2^{-n}/2^{-n} = (2^{-n} - 2^{-(n+1)})/2^{-n} \cong 1.)$$

All of this has a noble history going back more than half a century to the founding fathers, Lusin, Paley, Zygmund, Calderón and Stein.

The comparison of quadratic variation functionals and maximal functions via "good λ" inequalities was done twenty years ago (!) by Burkholder and me. These inequalities are the strongest comparisons I know. Because of this, their validity is limited to rather specific situations: we must have either a martingale structure (even here,we require special martingales like Haar series, or Brownian motion) or we must impose conditions on the convolution kernel φ. Until recently, we could only treat the cases where φ was the Poisson kernel (Burkholder and Gundy, 1973 - Studia Math.). The case where φ is the Gaussian kernel was done, partially, by Calderón and Torchinsky. (Advances in Math. circa 1975).

This brings us up to date, and to the problem addressed by Ileana Iribarren and me. Both the Poisson and Gaussian kernels are supported on the entire real line. For many approximation problems, it is desirable to consider kernels with compact support. Our goal was to find compactly supported convolution kernels for which we could prove "good λ" inequalites.

We did find a class of kernels that are compactly supported which do the job, provided we replace the Haar measure dt/t by $\Delta 2^{-n}/2^{-n} \cong 1$ at each integer. This is what happens in the Haar case. However, unlike the Haar case, we have full translation invariance for the sequence $\Phi_t(f)(x)$. That is,

$$\Phi_t(f)(x) = \int f(y)\varphi_t(x-y)dy$$

where $t = 2^{-n}, n \in \mathbf{Z}$. Basic examples of functions $\varphi(x)$ that provide us with approximation procedures that have "good λ" inequalities are the cardinal spline functions:

$$\varphi_2(x) = \begin{cases} 1-|x| & \text{if } |x| \leq 1 \\ 0 & \text{otherwise} \end{cases}$$

$$\varphi_3^3(x) := \int_{-\infty}^{\infty} \chi_{[-1/2,1/2]}(y)\varphi_2^2(x-y)dy$$

and

$$\varphi_n^n(x) := \int_{-\infty}^{\infty} \chi_{[1/2,1/2]}(y)\varphi_{n-1}^{n-1}(x-y)dy.$$

The class of densities for which we can prove the "good λ" inequalities may be specified in terms of the random variables that have the density in question. Let T be a symmetric, bounded random variable, that is, $T \cong -T$ where $\cong$ means "equal in distribution." Then T generates another variable Z by the dilation equation $2Z \cong T + Z'$, where T and Z' are supposed to be independent, and Z' has the same distribution as Z. Thus, $Z \cong \sum_{k=1}^{\infty} T_k/2^k$ where T_k is a sequence of independent random variables, each having the same distribution as T. The variable Z is assumed to have density φ. When T is binomial with probabilities and values are as follows, we obtain the first spline density, the triangular density:

$$T = \begin{cases} -1/2 & \text{with probability} \quad 1/4 \\ 0 & \text{with probability} \quad 1/2 \\ 1/2 & \text{with probability} \quad 1/4 \end{cases}$$

Thus, the triangular density is written as an average of dilated triangular densities as follows:

$$\varphi_0(x) = 1/4\varphi_1(x + 1/2) + 1/2\varphi_1(x) + 1/4\varphi_1(x - 1/2).$$

The definition of the quadratic functional associated with φ is given by a sum of terms, each of which is of the form

$$Q_n(x) = \mathcal{E}_t(\Phi_{n+1}^2(f)(x + t)) - (\Phi_n(f)(x))^2.$$

The functions $Q_n(x)$ are nonnegative since $\Phi_n(f)(x) = \mathcal{E}_t\Phi_{n+1}(f)(x + t)$. Here $\mathcal{E}_t$ is the average with respect to the distribution of $T/2^{n+1}$.

Here, my oral lecture finished. There is much more to tell, but perhaps next time.

Thank you for a wonderful visit. Dick Gundy.

CONFERENCISTAS

Carlos Benítez : Licenciado en Ciencias Matemáticas, Universidad Complutense de Madrid, 1965. Doctor en Ciencias Matemáticas, Universidad de Santiago de Compostela, 1971. Profesor de la Universidad de Extremadura, España.

El Profesor Benítez visitó el Departamento de Matemáticas de la Facultad de Ciencias de la Universidad de Los Andes, invitado por la Profesora Luisa Sánchez del Grupo de Análisis Funcional. Durante su permanecia, Septiembre-Octubre del 93, dictó un ciclo de conferencias sobre cuestiones relacionadas con la Geometría de los Espacios Normados.

Richard Gundy : Bachelor of Arts, Illionis College, 1955. Doctor of Philosophy, Indiana University, 1960 (Experimental Psychology). Doctor of Philosophy, University of Chicago, 1966 (Statistics). Profesor de Estadística y Matemáticas de la Universidad de Rutgers, Estados Unidos.

El Profesor Gundy participó en la Conferencia Internacional sobre Análisis Armónico y teoría de Operadores celebrada en Caracas del 3 al 8 de Enero de 1994, para festejar el octagésimo aniversario del Profesor Mischa Cotlar. Visitó los días 27 y 28 de Enero de 1994, el Departamento de Matemáticas de la Facultad de Ciencias de la Universidad de Los Andes invitado por el Profesor Víctor Padrón coordinador del Posgrado de Matemáticas.

Martin Jones : Bachelor of Arts, Warren Wilson College, 1979. Master of Sciences, University of South of Caroline, 1983. Doctor of Philosophy, George Institut of Technology, 1989. Profesor del Departamento de Matemáticas de la Universidad de Charleston, Estados Unidos.

El Profesor Jones visitó, del 24 al 28 de Mayo de 1993, el Departamento de Matemáticas de la Facultad de Ciencias de la ULA invitado por el Grupo de Matemáticas Aplicadas. En la actualidad el Profesor Jones se encuentra visitando nuestro Departamento bajo los auspicios de un beca Fulbright, su permanencia será hasta Febrero del 95.

Jorge Vargas : Licenciado en Matemática, Universidad Nacional de Córdoba, 1972. Doctor of Philosophy, Columbia University, 1977. Profesor de la Universidad Nacional de Córdoba, Argentina.

El Profesor Vargas visitó el Departamento de Matemáticas de la Facultad de Ciencias de la Universidad de Los Andes invitado por el Profesor Oswaldo Araujo del Grupo de Algebra. Durante su permanencia, Noviembre-Diciembre del 92, el Prof.. Vargas, dictó el curso Estructuras de módulos sobre álgebras de Lie semi-simple.

Wilfredo Urbina : Licenciado en Matemáticas, Universidad Central de Venezuela, 1978. Maestría en Matemáticas, Universidad Central de Venezuela, 1983. Doctor of Philosophy, Universidad de Minnesota, 1988. Profesor del Departamento de Matemáticas de la Facultad de Ciencias de la Universidad Central de Venezuela.

El profesor Urbina visitó, los días 29 y 30 de Julio de 1993, el Departamento de Matemáticas de la Facultad de Ciencias de la Universidad de Los Andes invitado por el Profesor Víctor Padrón coordinador del Posgrado de Matemáticas