

J. Albella Martínez  
A. Fernández Tojo  
E. García Portugués  
L. Búa Devesa

EDITORES

# As matemáticas do veciño

Actas do Seminario de Iniciación á Investigación

2012  
2013



INSTITUTO DE MATEMÁTICAS



**ACTAS DO SEMINARIO**

**DE**

**INICIACIÓN Á INVESTIGACIÓN**

**CURSO 2012 – 2013**

**Editores:**

Jorge Albella Martínez

Lucía Búa Devesa

Adrián Fernández Tojo

Eduardo García Portugués

© 2013 Seminario de Iniciación á Investigación.  
Instituto de Matemáticas da Universidade de Santiago de Compostela

**Coordina:**

Seminario de Iniciación á Investigación (SII)  
*seminarios3c@gmail.com*

**Edita:**

Instituto de Matemáticas da Universidade de Santiago de Compostela

**Imprime:**

DISERGAL S. Coop. Galega  
C/ Panasqueira, 1-5  
15895 Milladoiro - Ames  
A Coruña

**ISSN:** 2171-6536

**Depósito Legal:** C 2352-2013

*If you are a researcher, you are trying to figure out what the question is as well as what the answer is.*

Edward Witten (1951–).

*If we knew what it was we were doing, it wouldn't be called "research", would it?*

Albert Einstein (1879–1955).



# Prefacio

*Quen está aberto a un novo saber presta atención ás diferentes persoas. Disponse a escoitalas e a transformar as diferenzas en aprendizaxe.*

Icami Tiba (São Paulo, Brasil 2010).

Escoitar aos participantes do Seminario de Iniciación á Investigación (SII) sempre é para min unha experiencia de aprendizaxe. Primeiro por sentarme nunha cadeira da Facultade, como a estudante que entrou vai moitos anos, e gozar da “*aprendizaxe científica, un proceso sempre por rematar, que comenza cos nosos mestres, segue cos nosos compañeiros e remata cos nosos estudantes*” (Alberto Cabada, 2013).

Os prólogos das edicións do SII estanse a converter nunha gran mandala, que ao igual que o seu contido, constrúense a partir dun círculo que representa a unidade no coñecemento. Rosa Crujeiras pintou no prólogo do 2012 un ciclo de citas desta mandala ao que quixen sumar a do meu compañeiro Alberto. En cada nova capa, aos que nos honrades con esta encomenda, aportamos un pequeno incremento ao seu diámetro e, lendo os sentidos prólogos de ilustres predecesores, tratamos de redondear os traballos presentados por unha nova xeración que está a sementar as matemáticas do futuro.

Admiro e agradezo fondamente ás persoas que fixeron nacer o SII, e ás que o fan medrar curso a curso compartindo achegas e inqedanzas nas exposicións e nos debates. O nome das actas implica creatividade ao convidar a compartir sal e zucre matemática cos vecinos, transformando presuntas diferenzas de sabores en aprendizaxe, para conformar o mellor menú científico.

Os meus parabéns e ánimos a todas as persoas que dan valor e sabor a esta edición número nove: ás que asinan os traballos, ás que as convidaron, editaron e presentaron. Desexo que as agrias incertezas do momento actual sexan só unha inflexión nas vosas fructíferas traxectorias científicas, que agardo vos deixen por xustiza unha pegada de dulce felicidade.

Santiago de Compostela, 17 de setembro de 2013.

M. Elena Vázquez Cendón.



# Índice xeral

|   |           |
|---|-----------|
| <b>Introdución</b>  | <b>1</b>  |
| <b>Tania Pernas Castaño</b><br>“Las ecuaciones de Navier–Stokes”  | <b>3</b>  |
| <b>Yago Bea Besada</b><br>“Supermatemáticas”  | <b>9</b>  |
| <b>Ángel Manuel González Rueda</b><br>“Nuevas aplicaciones del valor de Shapley en juegos con estructura coalicional”                             | <b>15</b> |
| <b>Cristina Núñez García</b><br>“Modelos matemáticos para describir el comportamiento de una disolución de surfactante en la interfase aire–agua” | <b>21</b> |
| <b>María Oliveira</b><br>“Exploring circular data using CircSiZer”  | <b>27</b> |
| <b>Javier Seoane Bascoy</b><br>“Solitones de Yamabe en grupos de Lie Lorentzianos de dimensión tres”  | <b>33</b> |
| <b>F. Adrián Fernández Tojo</b><br>“La suma de todos los naturales y otras series divergentes”  | <b>39</b> |
| <b>Santiago Codesido Sánchez</b><br>“¿Por qué giran los electrones?”  | <b>45</b> |
| <b>Óscar A. Otero Zorraquiños</b><br>“Ecuaciones diferenciales de orden fraccionario”   | <b>51</b> |
| <b>Eduardo García Portugués</b><br>“¿Cómo contrastar la independencia estadística?”   | <b>57</b> |
| <b>Nizomjon Jumaniyazov</b><br>“On an inverse spectral problem for the Sturm–Liouville operator in a finite interval”                             | <b>63</b> |



# Introdución

O presente volume contén os resumos das charlas que se impartiron ao longo do ano 2013 no Seminario de Iniciación á Investigación (SII). Dito seminario, organizado por alumnos de doutoramento, ten lugar na Facultade de Matemáticas da Universidade de Santiago de Compostela e se encadra dentro das actividades do Instituto de Matemáticas.

O SII ten a súa orixe a comezos do ano 2005, como unha iniciativa dos alumnos de Terceiro Ciclo da Facultade e como resposta ás necesidades de crear un seminario que cumprise, cando menos, os seguintes obxectivos:

1. Fomentar o intercambio de coñecemento.
2. Proporcionar un lugar onde dar a coñecer os campos nos que cada un centra as súas investigacións.
3. Facilitar a práctica de falar en público, máis en concreto dar charlas e afacerse a escoitar e participar activamente neste tipo de eventos.
4. Proporcionar un marco onde se poidan levar a cabo as actividades necesarias para que cada quen saiba explicar as ideas fundamentais dos seus traballos incluso a persoas non especialistas no seu campo.

Por noveno ano consecutivo o SII acadou estes obxectivos básicos e ademais proporcionou un marco de intercambio de coñecemento entre alumnos dos distintos departamentos da Facultade. As charlas desenvolvéronse, salvo algunhas excepcións, de forma quincenal dende marzo ata xullo.

No referente á organización do SII, en 2013 tivo un novo comité organizador, formado por catro estudantes de doutoramento, que se encargaron tanto da coordinación do evento en si: calendario de charlas, anuncio das mesmas, reserva de aula, proporcionar o material necesario ao poñente, etc.; como da publicación deste anuario, onde se recolle un resumo de cada unha das charlas impartidas. Este mesmo comité organizador encargouse da confección deste volume e figura nel como comité editorial. Ademais é importante salientar que cada un dos resumos aquí recollidos pasou un proceso de revisión por parte dun alumno de Terceiro Ciclo, polo xeral dun departamento distinto ao do autor, co obxectivo de que así os resumos sexan comprensibles por aqueles que non son expertos no campo correspondente.

## Agradecementos

Quixeramos mencionar neste apartado que a organización do seminario tería sido, sen dúbida, moito máis difícil de non contarmos coa colaboración desinteresada de moita xente.

En primeiro lugar, os membros do comité organizador do SII deste último ano queremos agradecer aos membros do comité do ano anterior os consellos que nos prestaron, o empuxe para seguir organizando este seminario e o legado que nos deixaron.

Desexamos tamén dar as grazas a todos os que participaron no SII como oíntes, e moi especialmente aos poñentes e aos compañeiros que participaron no proceso de arbitraje: Yago Bea, Santiago Codesido, Ángel González, Nizomjon Jumaniyazov, Cristina Núñez, María Oliveira, Óscar Otero, Tania Pernas e Javier Seoane.

Agradecemoslle de xeito moi especial a Elena Vázquez Cendón a súa desinteresada colaboración na elaboración destas actas mediante a redacción do prefacio.

Santiago de Compostela, setembro de 2013.

O Comité Editorial.

## Las ecuaciones de Navier–Stokes

Tania Pernas Castaño

Instituto de Ciencias Matemáticas (CSIC, UAM, UC3M, UCM)

13 de marzo de 2013

### Reseña histórica

Se puede decir que la dinámica de fluidos comienza en el siglo XVII con el trabajo de Isaac Newton, que fue el primero en aplicar sus leyes de la mecánica a los movimientos de los fluidos. En el siglo siguiente, Leonhard Euler escribe por primera vez (1755) las ecuaciones diferenciales que rigen el movimiento de un fluido ideal, es decir, homogéneo, incompresible y sin viscosidad. Y finalmente, Navier (1822) y separadamente Stokes (1845), introdujeron en el modelo el término de viscosidad llegando a las ecuaciones que hoy denominamos *Navier–Stokes*.

### Ecuaciones de Euler y Navier–Stokes

Consideraremos que el fluido ocupa un dominio  $\Omega \subset \mathbb{R}^2$  o  $\mathbb{R}^3$ . En cada tiempo  $t$  las partículas del fluido tienen una correspondencia biyectiva con las coordenadas  $x = (x_1, x_2, x_3) \in \Omega$ . Caracterizamos el fluido por las siguientes funciones:

- Campo de velocidades,  $v(x, t) = (v_1(x, t), v_2(x, t), v_3(x, t))$ , que determina la velocidad que tiene una partícula en cada punto  $x \in \Omega$  del dominio y en cada tiempo  $t \in \mathbb{R}^+$ .
- Las presiones,  $p = p(x, t)$ , en el seno del fluido.
- La densidad,  $\rho = \rho(x, t)$ , del fluido.

Las ecuaciones pretenden modelar la evolución de un fluido a partir de la segunda ley de Newton, que asocia la aceleración de las partículas con las fuerzas que actúan sobre ellas, y con la ley de conservación de la masa.

Definimos la trayectoria de una partícula  $\alpha \in \Omega$  que está representada por  $X(\alpha, t)$  y satisface:

$$\begin{cases} \frac{\partial X(\alpha, t)}{\partial t} = v(X(\alpha, t), t), \\ X(\alpha, 0) = \alpha. \end{cases}$$

Si analizamos cómo cambia una función  $f$  según su trayectoria, tenemos la *derivada material*:

$$\frac{\partial f(X(\alpha, t), t)}{\partial t} = \frac{\partial f}{\partial t} + \frac{\partial X_1}{\partial t} \frac{\partial f}{\partial x_1} + \frac{\partial X_2}{\partial t} \frac{\partial f}{\partial x_2} + \frac{\partial X_3}{\partial t} \frac{\partial f}{\partial x_3} = \frac{\partial f}{\partial t} + (v \cdot \nabla) f \equiv \frac{Df}{Dt}.$$

Otra propiedad que se le va a pedir al fluido es incompresibilidad del mismo:

$$X(\alpha, t) \text{ es incompresible si } \text{vol}(\tilde{\Omega}) = \text{vol}(X(\tilde{\Omega}, t)) \text{ para todo } \tilde{\Omega} \subset \Omega,$$

que es equivalente a exigir al campo de velocidades que tenga divergencia nula:

$$X(\alpha, t) \text{ es incompresible} \iff \nabla \cdot v = 0.$$

Utilizando la segunda ley de Newton, tenemos que

$$\frac{D(\rho v)}{Dt} = \text{Fuerza},$$

por otro lado la ley de conservación de la masa junto con la incompresibilidad implica:

$$\frac{D\rho}{Dt} = 0.$$

Estas dos leyes dan lugar a las ecuaciones de Navier–Stokes:

$$\begin{cases} \rho \left( \frac{\partial v}{\partial t} + (v \cdot \nabla)v \right) = -\nabla p + \nu \Delta v + f, \\ \nabla \cdot v = 0, \\ \rho_t + (v \cdot \nabla)\rho = 0. \end{cases}$$

donde,  $\nu = \text{cte} \geq 0$  es la viscosidad del fluido y  $f = (f^1, f^2, f^3)$  una fuerza externa.

En el caso particular  $\nu = 0$  obtenemos las *ecuaciones de Euler*.

## Vorticidad, matriz de deformación y algunas soluciones exactas

Para un flujo incompresible, el campo de velocidades  $v(x, t)$  tiene su expansión en serie de Taylor para un punto fijo  $(x_0, t_0)$ :

$$v(x_0 + h, t_0) = v(x_0, t_0) + (\nabla v)(x_0, t_0)h + \mathcal{O}(h^2), \quad h \in \mathbb{R}^3.$$

La matriz  $\nabla v$  tiene una parte simétrica  $\mathcal{D}$  y otra antisimétrica  $\mathbf{\Omega}$ :

$$\mathcal{D} = \frac{1}{2}(\nabla v + \nabla v^t),$$

$$\mathbf{\Omega} = \frac{1}{2}(\nabla v - \nabla v^t).$$

$\mathcal{D}$  se llama matriz de deformación y  $\mathbf{\Omega}$  matriz de rotación.

Si el flujo es incompresible,  $\nabla \cdot v = 0$ , entonces  $\text{tr}\mathcal{D} = 0$ . Además, la vorticidad  $\omega$  del campo vectorial  $v$ ,

$$\omega = \text{curl}v = \left( \frac{\partial v_3}{\partial x_2} - \frac{\partial v_2}{\partial x_3}, \frac{\partial v_1}{\partial x_3} - \frac{\partial v_3}{\partial x_1}, \frac{\partial v_2}{\partial x_1} - \frac{\partial v_1}{\partial x_2} \right)$$

satisface

$$\mathbf{\Omega}h = \frac{1}{2}\omega \times h, \quad \forall h \in \mathbb{R}^3. \quad (1)$$

Utilizando estas definiciones podemos afirmar:

**Proposición 1** (Proposición 1.5 en [1]). *Sea  $\mathcal{D}(t)$  real, simétrica y con traza nula. Determinada la vorticidad por:*

$$\begin{cases} \frac{\partial \omega}{\partial t} = \mathcal{D}(t), \\ \omega(0) = \omega_0. \end{cases}$$

y la matriz antisimétrica tal que (1). Entonces

$$\begin{aligned} v(x, t) &= \frac{1}{2}\omega(t) \times x + \mathcal{D}(t)x, \\ p(x, t) &= -\frac{1}{2}[\mathcal{D}_t(t) + \mathcal{D}^2(t) + \mathbf{\Omega}^2(t)]x \cdot x, \end{aligned}$$

son soluciones de las ecuaciones de Euler y Navier–Stokes.

Veamos algunos ejemplos con esta proposición:

- Los *Jet Flows*. Son un ejemplos de soluciones estacionarias. Tomamos  $\omega_0 = 0$  y  $\mathcal{D} = \text{diag}(\gamma_1, \gamma_2, -(\gamma_1 + \gamma_2))$ ,  $\gamma_j > 0$  por (1):

$$\begin{aligned} v(x, t) &= (\gamma_1 x_1, \gamma_2 x_2, -(\gamma_1 + \gamma_2)x_3), \\ p(x, t) &= -\frac{\gamma_1}{2}x_1^2 - \frac{\gamma_2}{2}x_2^2 - \frac{\gamma_1 + \gamma_2}{2}x_3^2. \end{aligned}$$

Podemos observar que tanto la velocidad como la presión no dependen de  $t$ .

- Otro ejemplo son los *Strain Flows*. Ahora tomamos  $\omega_0 = 0$  y  $\mathcal{D} = \text{diag}(-\gamma, \gamma, 0)$ ,  $\gamma > 0$ . Así:

$$v(x, t) = (-\gamma x_1, \gamma x_2, 0)$$

Este flujo es irrotacional  $\omega = 0$  y forma un flujo plano independiente de la variable  $x_3$ .

- Un *vórtice*. Tomando  $\mathcal{D} = 0$  y  $\omega_0 = (0, 0, \omega_0)$ , tenemos:

$$v(x, t) = \left( -\frac{1}{2}\omega_0 x_2, \frac{1}{2}\omega_0 x_1, 0 \right).$$

Este flujo es un movimiento de rotación rígido en el plano  $x_1 - x_2$ .

## Cantidades globales conservadas

El dominio que consideramos es  $\Omega_n = \mathbb{R}^n$  (soluciones que decaen en el infinito) o  $\mathbb{T}^n$  (soluciones periódicas) con  $n = 2, 3$ . De las ecuaciones se deduce que se conservan las siguientes cantidades con fuerza  $f = 0$ :

- $\int_{\Omega_n} v \, dx$ , flujo total de velocidad;
- $\int_{\Omega_n} \omega \, dx$ , flujo total de vorticidad;
- $\int_{\Omega_n} v \cdot \omega \, dx$ , helicidad;
- $\int_{\Omega_n} |v|^2 \, dx$ , energía cinética con viscosidad nula (si  $\nu > 0$ , esta energía decae).

## Resultados destacados

- Existencia local (en tiempo) por el método de energía (ver [1]). Este método nos permite abordar las ecuaciones de Euler y Navier–Stokes simultáneamente, usando estimaciones que no dependen de la viscosidad  $\nu$ , y por tanto no cambian en el límite  $\nu \rightarrow 0$ . La existencia global está relacionada con la acumulación de la vorticidad.
- Existencia de soluciones débiles. En 1934 Leray [2] introdujo la noción de solución débil y probó su existencia para Navier–Stokes.
- Resultados de dato pequeño para  $n = 3$  y  $\nu > 0$  (ver [3]): Si la norma  $\|v_0\|_{H^{\frac{1}{2}}}$  (o la norma  $L^3$ ) es suficientemente pequeña, entonces existe solución global.
- Existencia global para datos iniciales particulares: Chemin, Gallagher y Paicu [4] prueban que en dimensión 3 (con  $\nu > 0$ ) existen soluciones globales en el tiempo con dato inicial que no es pequeño.

## Líneas de investigación

### El problema del Instituto Clay de Matemáticas

Consideramos un fluido viscoso, homogéneo e incompresible:

$$\begin{cases} \frac{\partial v}{\partial t} + (v \cdot \nabla)v = -\nabla p + \nu \Delta v + f, & \nu > 0, x \in \mathbb{R}^3, t \geq 0, \\ \nabla \cdot v = 0, \\ v(x, 0) = v_0. \end{cases}$$

- El dato inicial debe verificar las siguientes condiciones de regularidad:

$$|\partial_x^\alpha v_{0i}| \leq C_{\alpha,k}(1 + |x|)^{-k},$$

y fuerza exterior

$$|\partial_x^\alpha \partial_t^m f| \leq C_{\alpha,k,m}(1 + |x| + t)^{-k},$$

para todo  $\alpha, m, k > 0$ .

- Las soluciones admisibles son:
  - Para  $x \in \mathbb{R}^3$ ,  $(v, p) \in C^\infty(\mathbb{R}^3 \times [0, \infty))$  con decaimiento en el infinito de la presión y de energía finita.
  - Soluciones  $(v, p) \in C^\infty(\mathbb{T}^3 \times [0, \infty))$  periódicas y la presión de media cero.

### Problema del milenio:

- Sea  $v_0$  satisfaciendo las condiciones de regularidad. Demostrar que siempre existen soluciones admisibles.
- O encontrar  $v_0$  satisfaciendo las condiciones de regularidad y tal que no existe solución admisible con dato inicial  $v_0$ .

Para más detalles véase [5].

### Otras ecuaciones

Como acabamos de ver, uno de los problemas del milenio que queda por resolver es la existencia y unicidad global de solución clásica para la ecuación de Navier–Stokes tridimensional. Pues bien, relacionado con este problema están la existencia y unicidad de otras ecuaciones como las siguientes:

- **Las ecuaciones de Boussinesq:**

$$\begin{cases} \frac{\partial v}{\partial t} + (v \cdot \nabla)v = -\nabla p + (0, \theta), \\ \nabla \cdot v = 0, \\ \frac{\partial \theta}{\partial t} + (v \cdot \nabla)\theta = 0. \end{cases}$$

siendo unas ecuaciones bidimensionales tales que  $\theta$  es la temperatura del fluido.

- **Ecuaciones de fluidos en medios porosos:**

$$\begin{cases} \frac{\mu}{\kappa} = -\nabla p - (0, \rho)g, \\ \nabla \cdot v = 0, \\ \frac{\partial \rho}{\partial t} + (v \cdot \nabla)\rho = 0, \end{cases}$$

donde  $g$  es la gravedad. Ver [6].

▪ **Ecuaciones de la magnetohidrodinámica:**

$$\begin{cases} v_t + (v \cdot \nabla)v = -\nabla p + b \cdot \nabla b - \nu \Delta v, \\ b_t + (v \cdot \nabla)b = b \cdot \nabla v - \kappa \Delta b, \\ \nabla \cdot v = \nabla \cdot b = 0, \end{cases}$$

siendo  $b$  el campo magnético.

▪ **Ecuación Quasi-geostrófica:**

$$\begin{cases} \frac{\partial \theta}{\partial t} + (v \cdot \nabla)\theta = 0, \\ v = \nabla^\perp \psi, \end{cases}$$

con  $\nabla^\perp \psi = \left( \frac{\partial \psi}{\partial x_2}, -\frac{\partial \psi}{\partial x_1} \right)$  (ver [7]). Usando la vorticidad, estas ecuaciones son similares a la formulación de las ecuaciones de Euler:

$$\begin{cases} \frac{\partial \omega}{\partial t} + (v \cdot \nabla)\omega = 0, \\ v = \nabla^\perp \Delta^{-1} \omega. \end{cases}$$

## Bibliografía

- [1] Majda, A. y Bertozzi, A. (2002). *Vorticity and incompressible flow*, Cambridge University Press.
- [2] Leray, J. (1933). *Etude de diverses équations intégrales non linéaires et de quelques problèmes que pose l'hydrodynamique*, Journal de Mathématiques Pures et Appliquées, **12**, pp. 1–82.
- [3] Leray, J. (1934). *Sur le mouvement d'un fluide visqueux emplissant l'espace*, Acta Mathematica, **63**, pp. 193–248.
- [4] Chemin, J. Y., Gallagher, I. y Paicu, M. (2011). *Global regularity for some classes of large solutions to the Navier–Stokes equations*, Annals of Mathematics, **173**(2), pp. 983–1012.
- [5] Fefferman, C. *Clay Mathematics Institute, Millenium problems*.
- [6] Córdoba, D., Gancedo, F. y Orive, R. (2007). *Analytical behavior of 2D incompressible flow in porous media*, Journal of Mathematical Physics, **48**(6).
- [7] Constantin, P., Majda, A. y Tabak, E. (1994). *Formation of strong fronts in the 2D quasi-geostrophic thermal active scalar*, Nonlinearity, **7**, pp. 1495–1533.

## Supermatemáticas

Yago Bea Besada

Departamento de Física de Partículas

20 de marzo de 2013

### Introducción

A comienzos de los años 70 tiene lugar el nacimiento de la supersimetría (SUSY) [1, 2] en el ámbito teórico de la física de partículas elementales. Se trata de una propuesta que, aunque sin apoyo experimental hasta la fecha, es capaz de resolver algunos de los problemas más importantes de esta rama de la física, como el problema jerárquico del modelo estándar.

Junto a la supersimetría se han desarrollado las matemáticas sobre las que ésta se asienta, que han pasado a conocerse como supermatemáticas [3, 4, 5]. Se trata de una pequeña modificación de la geometría diferencial ordinaria, donde nos encontramos con objetos análogos: superespacios vectoriales, supervariedades, superálgebras de Lie, etc. Las supermatemáticas han mostrado poseer interés propio con independencia de su utilización en la supersimetría, encontrando aplicaciones en áreas como la topología o la geometría algebraica.

### Origen de las supermatemáticas: física de partículas y supersimetría

A principios de los años 70 se descubre la supersimetría dentro del marco de la física de partículas elementales. Se trata de una formulación teórica que propone que en la naturaleza las partículas aparecen por pares bosón-fermión. Bosones son partículas de espín entero  $(0, 1, 2, \dots)$  que obedecen a la estadística de Bose-Einstein, y fermiones son partículas de espín semientero  $(1/2, 3/2, \dots)$  que obedecen a la estadística de Fermi-Dirac. Todas las partículas elementales del modelo estándar se clasifican dentro de uno de estos dos grupos. Por ejemplo, el electrón es un fermión, así como los quarks que forman el protón. Los fotones, cuantos que forman la luz, son bosones. La propuesta de la supersimetría consiste en que por cada fermión del modelo estándar existe un correspondiente bosón y por cada bosón, un correspondiente fermión, ambos de la misma masa. Por ejemplo, el electrón debería tener una partícula compañera de espín 0 y misma masa, llamada selectrón. No obstante, en la

---

PALABRAS CLAVE: supermatemáticas; superespacios vectoriales; supervariedades; superálgebras de Lie.

naturaleza no se observan tales partículas. Por ello, se propone que la supersimetría está rota, y que la masa de estas partículas compañeras es mucho más elevada que la masa de las partículas del modelo estándar.

Sin embargo, no se ha obtenido hasta la actualidad ninguna evidencia experimental de la existencia de la supersimetría. ¿Por qué creer entonces en ella? El motivo fundamental por el cual la supersimetría ha despertado tanto interés es que resuelve de manera natural el problema jerárquico presente en el modelo estándar de la física de partículas. Este problema consiste en que la masa del bosón de Higgs debería sufrir una divergencia salvo que se ajusten de manera adecuada ciertos parámetros del modelo estándar, algo que para muchos teóricos es muy poco natural. Otro motivo importante es que también resuelve el problema de la unificación de las constantes de acople del modelo estándar a altas energías. Además, propone una partícula candidata a la materia oscura, el neutralino.

Por otra parte, la supersimetría goza de gran interés por ser un ingrediente esencial en la teoría de cuerdas, la gran candidata a ser la teoría fundamental de la naturaleza.

Vemos, por tanto, que los motivos para creer en la supersimetría son puramente teóricos.

## ¿Qué son las supermatemáticas?

El desarrollo de la supersimetría ha requerido de unas matemáticas propias sobre las que asentarse, que han pasado a conocerse como supermatemáticas. Éstas consisten en una ligera modificación de la geometría diferencial ordinaria, donde la conmutatividad se ve sustituida por la superconmutatividad: algunos elementos conmutan y otros anticonmutan. En términos físicos la conmutatividad se corresponde con elementos bosónicos, y la anticonmutatividad con elementos fermiónicos. Los objetos presentes en las matemáticas habituales tienen su análogo en las supermatemáticas. De este modo hablaremos de superespacios vectoriales, supermatrices, superdeterminantes, superálgebras conmutativas, supervariedades, superálgebras de Lie, etc. Las demostraciones de los teoremas en supermatemáticas encuentran gran similitud con las demostraciones de los teoremas clásicos, de tal manera que gran cantidad de éstos seguirán siendo válidos en las supermatemáticas. En particular, todas las construcciones presentes en las supermatemáticas contienen a las construcciones habituales. Por ejemplo, un superálgebra de Lie contiene a un álgebra de Lie en su estructura. Así, las supermatemáticas se pueden pensar como una generalización de las matemáticas habituales.

## Algunas definiciones

A continuación se muestran algunas definiciones que ilustran las ideas fundamentales que subyacen en las supermatemáticas.

**Definición 1.** Un **superespacio vectorial**  $V$  es un espacio vectorial  $\mathbb{Z}_2$ -graduado, es decir, que se puede expresar como suma directa de dos subespacios:

$$V = V_0 \oplus V_1,$$

junto con una aplicación denominada *paridad*:

$$\alpha : V_0 \cup V_1 \rightarrow \{0, 1\}, \quad \alpha(x) = 0 \text{ si } x \in V_0, \quad \alpha(x) = 1 \text{ si } x \in V_1.$$

Se denominan *elementos homogéneos* a los elementos de  $V_0$  y de  $V_1$ . En particular, los elementos de  $V_0$  se denominan *pares* y los elementos de  $V_1$  *impares*. Las propiedades sobre superespacios vectoriales se definen para los elementos homogéneos y se extienden a los demás por linealidad. La dimensión del superespacio vectorial es, por definición, el par  $(p, q)$ , donde  $p$  es la dimensión de  $V_0$  y  $q$  la dimensión de  $V_1$ , y se denota  $\dim(V) = p|q$ .

**Ejemplo 2.**  $\mathbb{R}^{p|q} = \mathbb{R}^p \oplus \mathbb{R}^q$ .

**Definición 3.** Se denomina **aplicación lineal par (impar)** a aquella aplicación lineal entre dos superespacios vectoriales que lleva elementos pares en pares e impares en impares (respectivamente pares en impares e impares en pares).

Cualquier aplicación lineal entre superespacios vectoriales se expresa de manera natural mediante una matriz bloque  $2 \times 2$ :

$$\begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}.$$

Este tipo de matrices se pueden generalizar a matrices cuyas entradas sean elementos de un superálgebra superconmutativa (ver Definición 4), en cuyo caso podrán definirse la trasposición, la conjugación compleja, la supertraza o el superdeterminante.

**Definición 4.** Un **superálgebra** es un superespacio vectorial  $V$  junto con una aplicación bilineal:

$$\begin{aligned} V \times V &\rightarrow V \\ (a, b) &\longrightarrow ab, \end{aligned}$$

que verifica:

$$\alpha(ab) = \alpha(a) + \alpha(b).$$

Donde  $\alpha(a) + \alpha(b)$  debe entenderse módulo 2. La actuación de esta aplicación se entiende sobre los elementos homogéneos, extendiéndose a los demás por linealidad.

**Definición 5.** Un **superálgebra superconmutativa** es un superálgebra que verifica:

$$ab = (-1)^{\alpha(a)\alpha(b)}ba, \quad \forall a, b \in V_0 \cup V_1.$$

**Ejemplo 6.** *Álgebra de Grassmann.* En particular podemos tomar el álgebra de Grassmann  $\Lambda(\mathbb{R}^3) = (\Lambda^0(\mathbb{R}^3) \oplus \Lambda^2(\mathbb{R}^3)) \oplus (\Lambda^1(\mathbb{R}^3) \oplus \Lambda^3(\mathbb{R}^3))$ .

**Definición 7.** Un *superálgebra de Lie* es un superálgebra que verifica:

$$\begin{aligned} [x, y] &= -(-1)^{\alpha(x)\alpha(y)}[y, x], \\ [x, [y, z]](-1)^{\alpha(x)\alpha(z)} + [z, [x, y]](-1)^{\alpha(z)\alpha(y)} + [y, [z, x]](-1)^{\alpha(y)\alpha(x)} &= 0, \end{aligned}$$

$\forall x, y, z \in V_0 \cup V_1$ .

**Ejemplo 8.** *El superálgebra lineal general  $gl(p|q)$  de las matrices  $(p|q) \times (p|q)$  con elementos en  $\mathbb{R}$  con el corchete definido del siguiente modo:*

$$[x, y] = xy - (-1)^{\alpha(x)\alpha(y)}yx.$$

Por último, puede mencionarse que las supermatemáticas se pueden generalizar. Por ejemplo, podríamos tomar espacios vectoriales gradados por un grupo arbitrario  $G$  (en lugar de  $\mathbb{Z}_2$ ).

## Resultados y aplicaciones

Muchos resultados de las matemáticas clásicas tienen su análogo en las supermatemáticas. Por ejemplo, de manera similar a la clasificación de las álgebras de Lie simples, existe una clasificación de las superálgebras de Lie simples.

Las supermatemáticas han encontrado importantes aplicaciones, entre las que destacan los resultados de dos Medalla Fields. Por una parte, Edward Witten (Medalla Fields en 1990) aplica la supergeometría en el ámbito de la teoría de Morse y las formas diferenciales. Por ejemplo, en [6] muestra que las desigualdades de Morse pueden obtenerse a partir de cierto hamiltoniano supersimétrico en mecánica cuántica. Por otra parte, M. L. Kontsevich (Medalla Fields en 1998) aplica la supergeometría en el ámbito de la teoría de deformaciones y a la cuantización de las variedades de Poisson.

Además, se pueden destacar algunos resultados adicionales de la supersimetría: la simetría especular, que consiste en una relación que se puede dar entre dos variedades de Calabi–Yau, la dualidad de Seiberg–Witten, la dualidad de Donaldson–Witten, etc.

## Bibliografía

- [1] Signer, A. (2009). *ABC of SUSY*, arXiv: 0905.4630.
- [2] Sohnius, M. F. (1985). *Introducing Supersymmetry*, Physics Reports **128**(2–3), pp. 39–204.
- [3] Berezin, F. A. (1987). *Introduction to Superanalysis*, Kluwer Academic Publishers, Mathematical Physics and Applied Mathematics.

- 
- [4] Varadarajan, V. S. (2004). *Supersymmetry for Mathematicians: An introduction*, American Mathematical Society, Courant Lecture Notes.
  - [5] Tuynman, G. M. (2004). *Supermanifolds and Supergroups: Basic Theory*, Springer.
  - [6] Witten, E. (1982). *Supersymmetry and Morse Theory*, Journal of Differential Geometry, **17**(4), pp. 661–692.



## Nuevas aplicaciones del valor de Shapley en juegos con estructura coalicional

Ángel Manuel González Rueda

Departamento de Estadística e Investigación Operativa

10 de abril de 2013

### Introducción

La Teoría de Juegos es una rama de las matemáticas que se dedica al estudio de los problemas de decisión en los que intervienen varios decisores. Podemos dividir los juegos en dos tipos: los juegos cooperativos y los no cooperativos. En los juegos cooperativos, que son los que trataremos en este trabajo, un conjunto de jugadores, que dispone de mecanismos para tomar acuerdos vinculantes, debe decidir cómo repartirse los beneficios de su cooperación. Cuando cualquier reparto es posible decimos que estamos ante un juego cooperativo con utilidad transferible (abreviadamente, juego TU). Por otro lado, en los juegos no cooperativos los jugadores no pueden llegar a acuerdos previos (por ejemplo, el dilema del prisionero).

Comenzaremos este trabajo con una serie de preliminares sobre los juegos cooperativos, los juegos con uniones a priori y la extensión multilineal de un juego. A continuación presentaremos un nuevo valor coalicional de Shapley que denotaremos por  $\Gamma$ . Llevaremos a cabo su caracterización axiomática, propondremos dos interpretaciones alternativas de este valor así como la forma de calcularlo a través de su extensión multilineal.

### Preliminares

**Definición 1.** *Un juego cooperativo con utilidad transferible (juego TU) es un par  $(N, v)$  donde  $N = \{1, \dots, n\}$  es el conjunto de jugadores y  $v : 2^N \rightarrow \mathbb{R}$  es una función que verifica que  $v(\emptyset) = 0$ . Denotaremos por  $G$  el conjunto de juegos TU.*

La función  $v$  se denomina función característica del juego. Dada una coalición  $S \subseteq N$ ,  $v(S)$  representa el pago que se pueden asegurar los jugadores de  $S$ , independientemente de cómo actúen el resto de los jugadores.

A los jugadores  $i, j \in N$  se les denomina *jugadores simétricos* si  $v(S \cup \{i\}) = v(S \cup \{j\})$  para todo  $S \subseteq N \setminus \{i, j\}$ .

**Definición 2.** *Un valor es una aplicación  $f$  que asigna a cada juego TU  $(N, v) \in G$  un vector  $f(N, v) \in \mathbb{R}^n$ , donde  $f_i(N, v)$  es el pago asignado al jugador  $i \in N$ .*

---

PALABRAS CLAVE: juego cooperativo; valor de Shapley; uniones a priori; valor coalicional.

Los dos valores más importantes relacionados con el concepto de juegos cooperativos son el valor de Shapley y el valor de Banzhaf.

**Definición 3** (Shapley, 1953). *El valor de Shapley,  $\varphi$ , asigna para cada juego  $(N, v) \in G$  y cada jugador  $i \in N$  el valor:*

$$\varphi_i(N, v) = \sum_{S \subseteq N \setminus \{i\}} \frac{s!(n-s-1)!}{n!} [v(S \cup \{i\}) - v(S)],$$

donde  $s = |S|$ .

**Definición 4** (Banzhaf, 1965). *El valor de Banzhaf,  $\beta$ , asigna para cada juego  $(N, v) \in G$  y cada jugador  $i \in N$  el valor:*

$$\beta_i(N, v) = \frac{1}{2^{n-1}} \sum_{S \subseteq N \setminus \{i\}} [v(S \cup \{i\}) - v(S)].$$

**Definición 5.** *Sea  $(N, v)$  un juego cooperativo de  $n$  jugadores. La extensión multilineal (EML) de  $v$  es la función real de  $n$  variables*

$$f(q_1, \dots, q_n) = \sum_{S \subseteq N} \prod_{j \in S} q_j \prod_{j \notin S} (1 - q_j) v(S).$$

**Definición 6.** *Consideremos un conjunto finito de jugadores  $N = \{1, \dots, n\}$ . Denotaremos el conjunto de todas las particiones de  $N$  por  $P(N)$ . Un sistema de uniones a priori o estructura de coaliciones  $P = \{P_1, \dots, P_m\}$  sobre  $N$  es una partición de  $N$  ( $P \in P(N)$ ), esto es,  $\bigcup_{k=1}^m P_k = N$  y  $P_k \cap P_h = \emptyset$  cuando  $k \neq h$ .*

Las estructuras de coaliciones triviales son  $P^n = \{\{1\}, \{2\}, \dots, \{n\}\}$ , donde cada unión es un jugador, y  $P^N = \{N\}$ , que es el sistema de uniones formado por la gran coalición.

Dado  $i \in N$ ,  $P(i)$  denota la familia de las estructuras de coaliciones sobre  $N$  donde  $\{i\}$  es una unión individual, es decir, si  $P \in P(i)$  entonces  $\{i\} \in P$ .

Si  $S \subseteq N$ , denotaremos por  $P^S$  el sistemas de uniones a priori cuyos miembros son  $S$  y todos los  $i$  tal que  $i \notin S$ . Notemos que si  $S = N$ ,  $P^S = P^N$  y si  $S = \{i\}$ ,  $P^S = P^n$ .

Sea  $P \in P(N)$ , para cada  $i \in P_k \in P$ ,  $P_{-i} \in P(i)$  denota la partición de  $N$  obtenida de  $P$  cuando el jugador  $i$  abandona su unión  $P_k$  y se aísla, i.e.,  $P_{-i} = \{P_h \in P : h \neq k\} \cup \{P_k \setminus \{i\}, \{i\}\}$ .

**Definición 7.** *Un juego TU con una estructura de coaliciones es una terna  $(N, v, P)$  donde  $(N, v) \in G(N)$  y  $P$  es una estructura de coaliciones sobre  $N$ .*

Denotaremos por  $G^{cs}$  el conjunto de todos los juegos TU con una estructura de coaliciones y por  $G^{cs}(N)$  el conjunto de juegos TU con una estructura de coaliciones y conjunto de jugadores  $N$ .

**Definición 8.** Dado  $(N, v, P) \in G^{cs}$  con  $P = \{P_1, \dots, P_m\}$ , tomemos  $i, j \in P_k \in P$  con  $i \neq j$ . Consideremos el conjunto  $\{i, j\}$  como un nuevo jugador  $p \notin N$  y denotemos por  $N^{ij} = (N \setminus \{i, j\}) \cup \{p\}$ . El  $\{ij\}$ -juego reducido  $(N^{ij}, v^{ij}, P^{ij}) \in G^{cs}$  de  $(N, v, P)$  se define para cada  $S \subseteq N^{ij}$  como  $v^{ij}(S) = v((S \setminus \{p\}) \cup \{i, j\})$  si  $p \in S$  y  $v^{ij}(S) = v(S)$  en otro caso. El sistema de uniones a priori  $P^{ij} = \{P_1^{ij}, \dots, P_m^{ij}\}$  se define como  $P_k^{ij} = (P_k \setminus \{i, j\}) \cup \{p\}$  y  $P_h^{ij} = P_h$  para  $h \neq k$ .

**Definición 9.** Por un valor coalicional en  $G^{cs}$  entendemos una función  $g$  que asigna un vector  $g(N, v, P) \in \mathbb{R}^n$  para cada juego con uniones a priori  $(N, v, P) \in G^{cs}(N)$ , donde  $g_i(N, v, P)$  es el pago del jugador  $i \in N$ .

Dado un valor  $f$  en  $G$ , un valor coalicional  $g$  en  $G^{cs}$  se dice que es un  $f$ -valor coalicional cuando  $g(N, v, P^n) = f(N, v)$ , para todo  $(N, v) \in G$ .

Uno de los valores más importantes que surgen en este ámbito es el valor de Banzhaf–Owen.

**Definición 10** (Owen, 1982). El valor de Banzhaf–Owen,  $\psi$ , es el valor de  $G^{cs}$  definido como

$$\psi_i(N, v, P) = \sum_{R \subseteq M \setminus \{k\}} \sum_{T \subseteq P_k \setminus \{i\}} \frac{1}{2^{m-1}} \frac{1}{2^{p_k-1}} \left[ v(Q \cup T \cup \{i\}) - v(Q \cup T) \right],$$

para todo  $i \in N$  y  $(N, v, P) \in G^{cs}$ , donde  $P_k \in P$  es la unión tal que  $i \in P_k$ ,  $M = \{1, \dots, m\}$ ,  $m = |M|$ ,  $p_k = |P_k|$  y  $Q = \bigcup_{h \in R} P_h$ .

El valor de Banzhaf–Owen es un valor coalicional de Banzhaf.

**Proposición 11.** Para todo  $(N, v, P) \in G^{cs}$  y todo  $i \in P_k$ ,  $P_k \in P$ ,

$$\Psi_i(N, v, P) = \beta_i \left( M \setminus \{k\} \cup P_k, v^{M \setminus \{k\} \cup P_k} \right),$$

donde

$$v^{M \setminus \{k\} \cup P_k}(R \cup T) = v \left( \bigcup_{h \in R} P_h \cup T \right),$$

para todo  $R \subseteq M \setminus \{k\}$  y  $T \subseteq P_k$ .

## El valor coalicional de Shapley $\Gamma$

El valor de Banzhaf–Owen asignaba a cada jugador el valor de Banzhaf de un cierto juego modificado. En este capítulo, proponemos un nuevo valor coalicional,  $\Gamma$ , que asigna a cada jugador el valor de Shapley del mismo juego modificado.

La definición formal del valor  $\Gamma$  es la siguiente:

**Definición 12.** El valor coalicional  $\Gamma$  se define como

$$\Gamma_i(N, v, P) = \sum_{R \subseteq M \setminus \{k\}} \sum_{T \subseteq P_k \setminus \{i\}} \frac{(r+t)!(m+p_k-r-t-2)!}{(m+p_k-1)!} [v(Q \cup T \cup \{i\}) - v(Q \cup T)],$$

para todo  $i \in N$  y  $(N, v, P) \in G^{cs}$ , donde  $P_k \in P$  es la unión tal que  $i \in P_k$ .

La siguiente proposición establece una posible interpretación para el valor  $\Gamma$ :

**Proposición 13.** Para todo  $(N, v, P) \in G^{cs}$  y cada jugador  $i \in P_k$ ,  $P_k \in P$ ,

$$\Gamma_i(N, v, P) = \varphi_i \left( M \setminus \{k\} \cup P_k, v^{M \setminus \{k\} \cup P_k} \right).$$

### Caracterización axiomática de $\Gamma$

Presentemos una serie de propiedades para un valor coalicional  $g$ .

- A1. Eficiencia.** Para todo  $(N, v) \in G$ ,  $\sum_{i \in N} g_i(N, v, P^n) = v(N)$ .
- A2. Simetría.** Si  $i, j \in N$  son jugadores simétricos en  $(N, v) \in G$ , entonces  $g_i(N, v, P^n) = g_j(N, v, P^n)$ .
- A3. Contribuciones marginales idénticas.** Si  $(N, v)$  y  $(N, w)$  son juegos TU con un conjunto de jugadores común  $N$ , y algún jugador  $i \in N$  satisface  $v(S \cup \{i\}) - v(S) = w(S \cup \{i\}) - w(S)$  para todo  $S \subseteq N \setminus \{i\}$ , entonces  $g_i(N, v, P^n) = g_i(N, w, P^n)$ .
- A4. Neutralidad ante abandonos individuales.** Si  $(N, v, P) \in G^{cs}$ ,  $P_k \in P$ , y  $i, j \in P_k$ ,  $i \neq j$ , entonces  $g_i(N, v, P) = g_i(N, v, P_{-j})$ .
- A5. Juego 1-cociente.** Si  $(N, v, P) \in G^{cs}$  y  $P \in P(i)$  para algún  $i \in N$ , entonces  $g_i(N, v, P) = g_k(M, v^P, P^m)$ , donde  $P_k = \{i\}$ .
- A6. Neutralidad para el juego reducido.** Si  $(N, v, P) \in G^{cs}$  y  $i, j \in P_k \in P$  con  $i \neq j$ , y  $l \in N \setminus P_k$ , entonces  $g_l(N, v, P) = g_l(N^{ij}, v^{ij}, P^{ij})$ , donde  $(N^{ij}, v^{ij}, P^{ij})$  es el  $\{ij\}$ -juego reducido de  $(N, v, P)$ .
- A7. Indiferencia para las uniones  $P^S$ .** Si  $(N, v, P) \in G^{cs}$  y  $P = P^S$  para algún  $S \subseteq N$ , entonces  $g_i(N, v, P) = g_i(N, v, P^n)$ , si  $i \in S$ .

Se verifica el siguiente resultado:

**Proposición 14.** Un valor coalicional  $g$  satisface A1, A2 y A3 si, y sólo si, es un valor coalicional de Shapley, i.e. si

$$g_i(N, v, P^n) = \varphi_i(N, v) \text{ para todo } i \in N \text{ y todo } (N, v) \in G.$$

A continuación proponemos dos caracterizaciones axiomáticas para el valor  $\Gamma$ :

**Proposición 15.** *El valor coalicional  $\Gamma$  satisface A1–A7.*

**Teorema 16.**  *$\Gamma$  es el único valor coalicional que satisface A1–A5. Equivalentemente,  $\Gamma$  es el único valor coalicional de Shapley que satisface A4 y A5.*

En el siguiente teorema, obtenemos una caracterización axiomática alternativa del valor propuesto, empleando A6 y A7 en lugar de A4 y A5.

**Teorema 17.**  *$\Gamma$  es el único valor coalicional que satisface A1–A3, A6, y A7. Equivalentemente,  $\Gamma$  es el único valor coalicional de Shapley que satisface A6 y A7.*

## Interpretaciones de $\Gamma$

### El “Valor de Shapley Modificado” de un juego

Consideremos un jugador  $i \in P_k$ , sea  $\pi$  una permutación de jugadores en  $N$  en la cual los jugadores en  $P_h$ ,  $h \neq k$ , actúan como un bloque, i.e., los jugadores de cada una de dichas uniones actúan como un único jugador, y  $\Pi^*(N)$  el conjunto de tales permutaciones. Denotaremos por  $B^\pi(i) = \{j \in N \text{ tal que } \pi(j) < \pi(i)\}$  el conjunto de jugadores que preceden a  $i$  con el orden dado por  $\pi$ . Suponemos que:

1. Los jugadores acuerdan acudir a un cierto punto de negociación.
2. Todos los posibles órdenes de llegada en los que los jugadores de  $P_h$ ,  $h \neq k$ , actúan como un bloque son equiprobables.
3. Cuando llega un jugador, recibe como pago su contribución marginal a la coalición formada por los jugadores que llegaron antes que él.

Entonces, si  $i \in P_k$ :

$$\Gamma_i(N, v, P) = \frac{1}{(m + p_k - 1)! \prod_{h=1, h \neq k}^m p_h!} \sum_{\pi \in \Pi^*(N)} [v(B^\pi(i) \cup \{i\}) - v(B^\pi(i))].$$

De este modo, el valor coalicional  $\Gamma$  aparece como un “Valor de Shapley Modificado” del juego original. Es decir,  $\Gamma_i$  es el valor de Shapley del jugador  $i$  cuando los jugadores de  $P_k$  llegan en cualquier orden (los órdenes de llegada para los jugadores de  $P_k$  son equiprobables), pero los jugadores pertenecientes a uniones distintas de  $P_k$  solo pueden actuar como un bloque.

### El Valor de Shapley de un juego modificado

El valor coalicional  $\Gamma$  admite la siguiente interpretación: consideremos las uniones distintas de  $P_k$  como votantes individuales (i.e, cada unión distinta de  $P_k$  va a ser un único jugador individual) y los votantes pertenecientes a la unión  $P_k$  como votantes independientes. De este modo, para cada votante en  $P_k$ , la situación puede ser descrita por el  $(M \setminus \{k\}) \cup P_k$ -juego

$$v^{M \setminus \{k\} \cup P_k}(R \cup T) = v \left( \bigcup_{h \in R} P_h \cup T \right),$$

para todo  $R \subseteq M \setminus \{k\}$  y todo  $T \subseteq P_k$ .

Entonces, tenemos que:

$$\Gamma_i(N, v, P) = \varphi_i(M \setminus \{k\} \cup P_k, v^{M \setminus \{k\} \cup P_k}), \forall i \in P_k.$$

Entonces el valor coalicional  $\Gamma$  para cualquier votante de  $P_k$  viene dado por el valor de Shapley del  $(M \setminus \{k\}) \cup P_k$ -juego descrito anteriormente. Notemos que el juego considerado para cada unión es diferente.  $\Gamma_i(N, v, P)$  puede interpretarse como la probabilidad de que el jugador  $i \in P_k$  sea pivot<sup>1</sup> asumiendo una distribución uniforme de las permutaciones de  $(M \setminus \{k\}) \cup P_k$ .

### Extensión multilinear de $\Gamma$

Una de las principales dificultades del cálculo de un valor con su definición es que requiere la computación de una gran cantidad de términos. De este modo, para calcular un valor de forma más sencilla, se suele recurrir con frecuencia a su extensión multilinear. Sea  $(N, v)$  un juego con estructura coalicional  $P = \{P_1, \dots, P_m\}$ , puede calcularse  $\Gamma_i$ ,  $i \in P_k$ , siguiendo los siguientes pasos:

1. Obtener la extensión multilinear  $f(q_1, \dots, q_n)$  del juego  $(N, v)$ .
2. Para todo  $h \neq k$  y para todo  $i \in P_h$ , reemplazar la variable  $q_i$  por  $p_h$ . Se obtiene así una nueva función de  $q_i$ ,  $i \in P_k$ , y  $p_h$ ,  $h \neq k$ .
3. En la función obtenida en el paso 2, reducir todos los exponentes a 1, i.e., reemplazar cada  $p_l^n$  con  $n \geq 1$ , por  $p_l$ . Con esto se ha obtenido una nueva función multilinear:

$$g((q_i)_{i \in P_k}, (p_h)_{h \neq k}).$$

4. Finalmente, para obtener el valor coalicional  $\Gamma$ , calcular:

$$\Gamma_i(N, v, P) = \int_0^1 \frac{\partial g}{\partial q_i}(t, \dots, t) dt.$$

### Bibliografía

- [1] Alonso–Mejjide, J. M., Casal–Méndez, B., Lorenzo–Freire, S., González–Rueda, A. M. (2013). *Axiomatic of the Shapley value of a game with a priori unions*, TOP, to appear.
- [2] Laruelle, A. y Valenciano, F. (2004). *On the meaning of Owen–Banzhaf coalitional value in voting situations*, Theory and Decision **56**, pp. 113–123.
- [3] Shapley, L. S. (1953). *A value for  $n$ -person games*, en *Contributions to the Theory of Games II*, H. W. Kuhn, A. W. Tucker (Eds.), Princeton University Press, pp. 307–317.

<sup>1</sup>Dada una coalición  $S \subseteq N \setminus \{i\}$ , se dice que el jugador  $i$  es pivot si  $S$  es una coalición perdedora y  $S \cup \{i\}$  es ganadora.

## Modelos matemáticos para describir el comportamiento de una disolución de surfactante en la interfase aire–agua

Cristina Núñez García

Departamento de Matemática Aplicada

24 de abril de 2013

Los surfactantes son compuestos químicos cuyas moléculas constan de dos partes diferentes: una parte hidrófila y una parte hidrófoba. Esta estructura marca el comportamiento de las moléculas de surfactante cuando éstas entran en contacto con agua. Cuando se forma una nueva superficie en una disolución acuosa de surfactante, las moléculas de este compuesto viajan a la superficie con el fin de alcanzar un estado termodinámico más estable. Este hecho causa un cambio en las propiedades de la superficie del agua, entre las cuales está la tensión superficial. La tensión superficial es una fuerza que actúa tangencialmente por unidad de longitud en la superficie de un líquido en equilibrio como resultado de las fuerzas de unión entre las moléculas superficiales del mismo. Al incorporarse a la superficie de la disolución, lo que hacen los surfactantes es romper las uniones entre las moléculas de agua, reduciendo drásticamente su tensión superficial. El análisis de la variación de la tensión superficial (tensión superficial dinámica) de una disolución de surfactante en la interfase aire–agua es un tema que se ha convertido en objeto de estudio por sus importantes aplicaciones en diversas áreas como la biología, la medicina, el procesado de alimentos, la metalurgia, etc (ver [1]). La duración de este proceso dinámico puede variar desde unos pocos segundos hasta varias horas e incluso días dependiendo de la temperatura, la salinidad, el tipo de surfactante y su concentración.

Como se puede intuir, el análisis de la tensión superficial dinámica está ampliamente ligado al estudio del transporte de moléculas en la disolución, que tiene lugar mediante dos mecanismos diferentes: la difusión y la adsorción/desorción. Para entender el proceso es importante tener en cuenta la zona llamada “subsuperficie”(ver Figura 1), una frontera imaginaria situada a unos pocos diámetros moleculares por debajo de la superficie, que separa la región donde solamente tiene lugar la difusión de aquella en la cual se lleva a cabo la adsorción/desorción. Para describir el proceso de adsorción/desorción existen dos modelos:

- *Modelo controlado por difusión*, en el que la escala de tiempo para alcanzar el equilibrio entre la superficie y la subsuperficie es mucho menor que la escala de

---

PALABRAS CLAVE: surfactante; difusión; adsorción/desorción; tensión superficial; método de elementos finitos.

tiempo necesaria para la difusión. En este caso, se supone que las moléculas de surfactante, una vez llegan a la subsuperficie, son adsorbidas directamente a la superficie. Además, en este modelo, la relación entre las concentraciones superficiales y subsuperficiales viene dada por una ecuación llamada *isoterma de adsorción*. Dos de las isotermas más estudiadas son la isoterma de Henry y la isoterma de Langmuir. La primera es la más simple y establece una dependencia lineal entre las concentraciones de la superficie y la subsuperficie, mientras que la segunda define una dependencia no lineal entre ambas concentraciones.

- *Modelo cinético mixto*. En este caso, las escalas de tiempo para los procesos de adsorción/desorción y de difusión son comparables. El equilibrio entre la superficie y la subsuperficie no se alcanza de manera instantánea, sino que las moléculas de surfactante tienen que superar una barrera energética para colocarse en una orientación correcta o para encontrar una posición vacante en la superficie. En este tipo de modelos, una expresión cinética relaciona el balance entre las tasas de cambio de la adsorción y desorción.

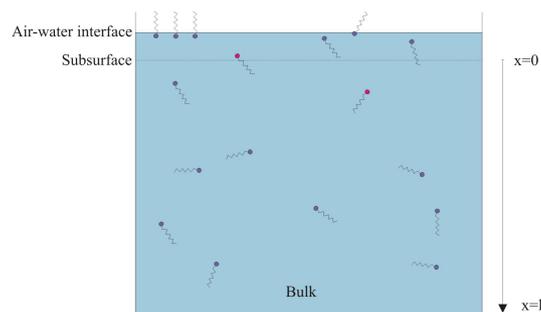


Figura 1: Subsuerficie e interfase aire-agua

Desde el punto de vista matemático, el proceso de difusión es modelado por la ecuación de difusión en una dimensión, junto con unas condiciones de contorno e iniciales adecuadas. Además, la dinámica de adsorción/desorción se describe a través del correspondiente modelo de adsorción, que se acopla al sistema a través de la condición de contorno en la subsuperficie, que involucra a la derivada temporal de la concentración superficial. Las incógnitas del problema son las concentraciones en la superficie y en el seno de la disolución. En este estudio, nos centramos en analizar el problema de difusión teniendo en cuenta el modelo cinético mixto lineal.

## Modelado matemático del problema

Denotemos por  $x$  la distancia desde la subsuperficie y por  $c(t, x)$  la concentración de surfactante en el instante  $t \in [0, T]$  y en el punto  $x \in [0, l]$ . La frontera  $x = 0$  del intervalo espacial corresponde a la localización de la subsuperficie. Denotando por

$\Gamma(t)$  la concentración superficial, que depende del tiempo, y teniendo en cuenta la ley de Fick, consideramos la ecuación de difusión:

$$\frac{\partial c}{\partial t}(t, x) - D \frac{\partial^2 c}{\partial x^2}(t, x) = 0, \quad t > 0, \quad x \in (0, l), \quad (1)$$

junto con las condiciones de contorno:

$$D \frac{\partial c}{\partial x}(t, 0) = \frac{d\Gamma}{dt}(t), \quad t > 0, \quad (2)$$

$$c(t, l) = c_b, \quad t > 0, \quad (3)$$

y las condiciones iniciales:

$$c(0, x) = c_0(x), \quad x \in (0, l), \quad (4)$$

$$\Gamma(0) = \Gamma_0. \quad (5)$$

En las ecuaciones previas,  $D$  denota el coeficiente de difusión y  $c_b$  es una constante positiva que representa la concentración en el seno de la disolución. Además,  $c_0(x)$  es una función definida en  $[0, l]$ , que es igual a  $c_b$  en  $x = l$ . Como la concentración superficial,  $\Gamma$ , es una incógnita del sistema, debemos de dar una condición adicional para cerrar el problema. Por eso, o bien el modelo controlado por difusión o bien el cinético mixto debe de ser considerado. En este trabajo, consideramos el segundo, teniendo en cuenta la expresión cinética más simple, que modela la transferencia de moléculas de surfactante entre la superficie y la subsuperficie mediante la siguiente ecuación diferencial ordinaria:

$$\frac{d\Gamma}{dt}(t) = k_H^a c(t, 0) - k_H^d \Gamma(t), \quad t > 0, \quad (6)$$

siendo  $k_H^a$  y  $k_H^d$  las constantes de adsorción y desorción, respectivamente.

Asumiendo regularidad suficiente, la ecuación (6), junto con la condición inicial (5), puede ser integrada para obtener:

$$\Gamma(t) = \Gamma_0 e^{-k_H^d t} + k_H^a e^{-k_H^d t} \int_0^t e^{k_H^d \tau} c(\tau, 0) d\tau.$$

Por lo tanto, la condición de contorno (2) quedaría:

$$D \frac{\partial c}{\partial x}(t, 0) = k_H^a c(t, 0) - k_H^d \left( \Gamma_0 e^{-k_H^d t} + k_H^a e^{-k_H^d t} \int_0^t e^{k_H^d \tau} c(\tau, 0) d\tau \right). \quad (7)$$

Teniendo en cuenta esta nueva condición de contorno, el problema que queremos analizar ahora es (1), (3), (4) junto con (7). Con el fin de simplificar los cálculos y sin pérdida de generalidad, supondremos que  $c_b$  es igual a cero y por lo tanto, imponemos una condición Dirichlet homogénea en el extremo derecho del intervalo espacial.

## Formulación variacional del problema

Antes de escribir la formulación variacional del problema, introducimos los espacios que usaremos más adelante. Sean  $H = L^2(0, l)$ ,  $V = \{v \in H^1(0, l); v(l) = 0\}$ ,  $W_2(0, T; V) = \{v \in L^2(0, T; V); \dot{v} \in L^2(0, T; V')\}$ , donde la derivada se entiende en el sentido de las distribuciones. Además, denotamos por  $\gamma_0 : H^1(0, l) \rightarrow \mathbb{R}$ , al operador traza en  $x = 0$ .

Para obtener la formulación débil del problema (1), (3), (4) y (7) multiplicamos la ecuación (1) por una función test definida en  $[0, l]$  y que valga cero en el extremo derecho del intervalo espacial, integramos en  $(0, l)$ , usamos la fórmula de integración por partes y la ecuación (7) y obtenemos el siguiente problema:

**Problema  $P_V$ .** Dado  $c_0 \in H$ , encontrar una función  $c \in W_2(0, T; V)$  tal que

$$\begin{aligned} \langle \dot{c}(t), v \rangle_{V' \times V} + D((c(t), v)) + k_H^a \gamma_0 c(t) \gamma_0 v = k_H^d \Gamma_0 e^{-k_H^d t} \gamma_0 v \\ + k_H^d k_H^a e^{-k_H^d t} \int_0^t e^{k_H^d \tau} \gamma_0 c(\tau) d\tau \gamma_0 v, \end{aligned}$$

para c.t.p.  $t \in (0, T)$ ,  $\forall v \in V$  y  $c(0) = c_0$ .

**Teorema 1.** Sean  $k_H^a$ ,  $k_H^d$  y  $D$  constantes positivas. Si  $c_0 \in H$ , entonces existe una única solución  $c \in W_2(0, T; V)$  del Problema  $P_V$ .

### Demostración:

La demostración de la existencia y unicidad de solución del Problema  $P_V$  está detallada en [2], y se basa en resultados clásicos para ecuaciones parabólicas lineales y técnicas de punto fijo.  $\square$

## Discretización del problema: análisis numérico

En esta sección consideramos una discretización en espacio y tiempo del Problema  $P_V$ . La discretización temporal se realiza teniendo en cuenta una combinación híbrida de los esquemas de Euler explícito e implícito, mientras que, para la discretización espacial, se utiliza el método de elementos finitos, gracias al cual se obtiene un espacio de dimensión finita  $V^h \subset V$ , que aproxima al espacio  $V$ .

Denotamos por  $h > 0$  el parámetro de discretización espacial,  $k > 0$  el tamaño de los elementos de la malla temporal, con nodos  $t_n = nk$ , para  $n = 0, \dots, N$ . Además, para una función continua  $f(t)$ , usamos la notación  $f_n = f(t_n)$  y para una sucesión  $\{f_n\}_{n=0}^N$  denotamos  $\delta f_n = (f_n - f_{n-1})/k$ . Teniendo en cuenta todo esto, se obtiene el problema discretizado que se escribe a continuación.

**Problema  $P^{hk}$ .** Encontrar  $c^{hk} = \{c_n^{hk}\}_{n=0}^N$  tal que  $c_0^{hk} = c_0^h$ , para  $n = 1, \dots, N$

y para todo  $v^h \in V^h$ ,

$$\begin{aligned} & (\delta c_n^{hk}, v^h)_H + D((c_n^{hk}, v^h)) + k_H^a \gamma_0(c_n^{hk}) \gamma_0(v^h) \\ &= \left( k_H^d \Gamma_0 e^{-k_H^d t_n} + k_H^d k_H^a k \sum_{j=0}^{n-1} e^{k_H^d (t_j - t_n)} \gamma_0(c_j^{hk}) \right) \gamma_0(v^h), \end{aligned}$$

siendo  $c_0^h \in V^h$  una aproximación adecuada de la condición inicial  $c_0$ .

En lo sucesivo, consideramos el siguiente espacio de elementos finitos:

$$V^h = \{v^h \in C([0, l]); v_{|[a_{i-1}, a_i]}^h \in P_1([a_{i-1}, a_i]), \forall i = 1, \dots, M, v^h(l) = 0\}.$$

**Teorema 2.** *Supongamos que se verifican las hipótesis del Teorema 1 y que  $c$  verifica las siguientes condiciones de regularidad:*

$$\begin{aligned} c &\in \mathcal{C}([0, T]; V) \cap \mathcal{C}([0, T]; H^2(0, l)) \cap \mathcal{C}^1([0, T]; H), \\ \dot{c} &\in L^2(0, T; V), \quad \ddot{c} \in \mathcal{C}([0, T]; H). \end{aligned}$$

Entonces, el algoritmo converge linealmente, es decir, existe una constante positiva  $\beta > 0$ , independiente de  $h$  y  $k$ , tal que

$$\max_{0 \leq n \leq N} \|c_n - c_n^{hk}\|_H \leq \beta(h + k).$$

La demostración se puede ver en [2].

## Ejemplo: simulación para el hexanol

Con el fin de mostrar el comportamiento de este modelo, consideramos una disolución de hexanol teniendo en cuenta los siguientes datos (ver [1]):

$$\begin{aligned} c_b &= 3,44 \text{ mol/m}^3, \quad D = 7,16 \times 10^{-10} \text{ m}^2/\text{s}, \quad l = 10^{-4} \text{ m}, \\ T &= 0,5 \text{ s}, \quad \Gamma_0 = 0 \text{ mol/m}^2, \quad k_H^a = 1,73 \times 10^{-4} \text{ m/s}, \quad k_H^d = 157 \text{ s}^{-1}. \end{aligned}$$

Además, suponemos que, en el instante inicial, la concentración de surfactante en toda la disolución es constante, siendo su valor igual a  $c_b$ . En la gráfica de la izquierda de la Figura 2 se representa la concentración de surfactante en el instante final que, como se puede observar, es casi uniforme en toda la disolución. En la gráfica de la derecha de esta misma figura se dibuja la evolución de la concentración de la subsuperficie a lo largo del tiempo. En un primer momento, la concentración de la subsuperficie desciende debido a que la superficie se encuentra vacía y las moléculas de surfactante adsorben directamente en ella. Sin embargo, a medida que pasa el tiempo y la superficie se va llenando, las moléculas de surfactante, que no encuentran un puesto vacante en la superficie, permanecen en la subsuperficie, aumentando así su concentración.

Una vez que la concentración superficial es conocida, se puede calcular la tensión superficial  $\sigma$  mediante la ecuación superficial de estado:

$$\sigma(t) = \sigma_0 - n R \theta \Gamma(t),$$

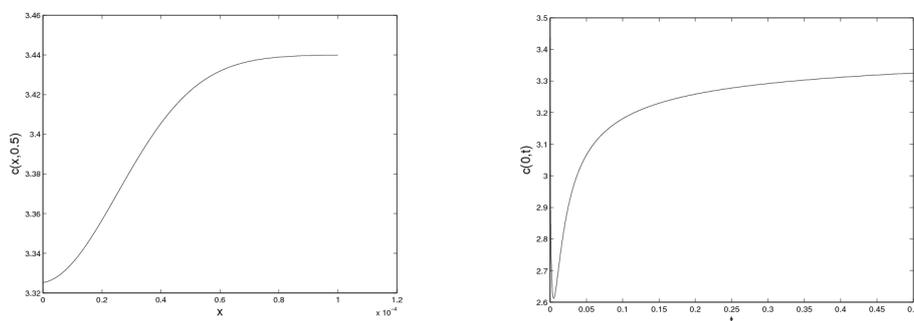


Figura 2: Concentración en el instante final (izquierda) y evolución en tiempo de la concentración en la subsuperficie (derecha). Resultados obtenidos con MATLAB.

donde  $\sigma_0 = 0,072$  N/m denota la tensión superficial del agua pura,  $\theta = 293$  K es la temperatura,  $R = 8,31$  J/(K mol) representa la constante de los gases y  $n$  es una constante que vale 1 para surfactantes no iónicos. En la Figura 3 se puede ver la evolución de la tensión superficial. En un primer momento, la tensión superficial de la disolución es igual a la del agua, pero, la incorporación de moléculas de surfactante en la superficie hace que esta tensión disminuya hasta el valor de equilibrio.

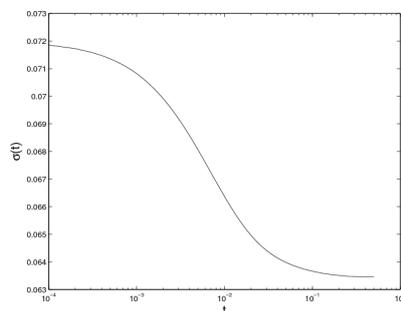


Figura 3: Evolución en tiempo de la tensión superficial, en escala logarítmica.

## Bibliografía

- [1] Chang, C. H. y Franses, E. I. (1995). *Adsorption dynamics of surfactants at the air/water interface: a critical review of mathematical models, data and mechanisms*, Colloids and Surfaces, **100**, pp. 1–45.
- [2] Fernández, J. R., Muñiz, M. C. y Núñez, C. (2012). *A mixed kinetic-diffusion surfactant model for the Henry isotherm*, Journal of Mathematical Analysis and Applications **389**, pp. 670–684.

## Exploring circular data using CircSiZer

María Oliveira

Department of Statistics and Operations Research

8th May, 2013

### Introduction

A circular observation can be defined as a point on a circle of unit radius, or a unit vector (i.e., a direction in the plane) and is usually expressed as an angle relative to some fixed reference point. Circular data appear naturally in a large variety of scientific fields such as meteorology, ecology, biology and medicine, among others. However, the periodic nature of this type of data sets apart circular statistical analysis from standard methods for linear data.

Both for linear and circular data, smoothing methods are useful for finding important and nonobvious structure in data. But, some of the features discovered in this way can be spurious sampling artifacts. Moreover, as in any nonparametric procedure, a smoothing parameter controlling the global aspect of the estimator and its dependence on the sample must be chosen. In order to provide a graphical tool to assess significance features of the underlying density without relying on a smoothing parameter, the SiZer (SIGNificant ZERo crossings of derivatives) map was proposed by [1] for linear data. This method is based on studying statistical significance of zero crossings of smoothed estimates.

In the context of circular data, circular kernel density estimators provide a route to visualize the data structure underlying a random phenomenon. Thus, based on this kind of smoothers, the SiZer approach can be fitted to the circular data setting yielding a visualization tool introduced by [4], namely CircSiZer, for the assessment of statistically significant features for this type of data.

The aim of this work is to introduce the CircSiZer method. For that purpose, a brief overview on kernel density estimation for circular data is provided. After that, the construction of CircSiZer map is introduced. Finally, the performance of CircSiZer is illustrated with some simulated and real data sets.

---

KEYWORDS: CircSiZer; circular data; nonparametric density estimation.

## Nonparametric circular kernel density estimation

Given a random sample of angles  $\Theta_1, \Theta_2, \dots, \Theta_n \in [0, 2\pi)$  from a density  $f$ , the circular kernel density estimator of  $f$  is defined as:

$$\hat{f}(\theta; \nu) = \frac{1}{n} \sum_{i=1}^n K_\nu(\theta - \Theta_i), \quad 0 \leq \theta < 2\pi,$$

where  $K_\nu$  is the circular kernel function with concentration parameter  $\nu > 0$  (see [2]). As a circular kernel, the von Mises distribution can be considered. The von Mises distribution,  $vM(\mu, \kappa)$ , is a symmetric unimodal distribution characterized by a mean direction  $\mu \in [0, 2\pi)$ , and a concentration parameter  $\kappa \geq 0$ , with probability density function

$$g(\theta; \mu, \kappa) = \frac{1}{2\pi I_0(\kappa)} \exp\{\kappa \cos(\theta - \mu)\}, \quad 0 \leq \theta < 2\pi,$$

where  $I_0$  denotes the modified Bessel function of order 0. With this specific kernel, the density estimator is given by:

$$\hat{f}(\theta; \nu) = \frac{1}{n(2\pi)I_0(\nu)} \sum_{i=1}^n \exp\{\nu \cos(\theta - \Theta_i)\}, \quad 0 \leq \theta < 2\pi, \quad (1)$$

which can be seen as a mixture of von Mises distributions centered in  $\Theta_i$  and with concentration parameter  $\nu$ .

A critical issue is the choice of the smoothing parameter  $\nu$  for (1), with large values leading to highly variable (undersmoothed) estimators and small values showing an opposite behaviour (see Figure 1).

## CircSiZer map

Apart from the lack of a uniformly superior rule for selecting the smoothing parameter, from a practical point of view, the exploration of the estimators at different smoothing levels (for a range of reasonable values of the smoothing parameter, between oversmoothing and undersmoothing levels) provides more in-depth information about the available data. However, significant features in the underlying data structure should be effectively disentangled from sampling artifacts. Features like peaks and valleys of a smooth curve can be characterized in terms of zero crossings of derivatives. Hence, the significance of such features can be judged from statistical significance of zero crossings or equivalently by the sign changes of derivatives. This idea has been successfully exploited by [1] in developing a simple yet effective tool called SiZer for exploring significant structures in density and regression curves. Taking into account the circular nature of data, SiZer technique is developed for circular data by [4], namely CircSiZer.

In the usual inferential approach in the statistical literature, the spotlight is placed on the true underlying curve  $f$  and doing inference on it, in particular, based on confidence bands. A crucial problem in nonparametric estimation is that  $f(\theta; \nu) = E(\hat{f}(\theta; \nu))$  is not necessarily equal to  $f(\theta)$ , involving an inherent bias specially for small values of  $\nu$  (see Figure 1, left). The bias can be reduced by taking large values of  $\nu$ , but in this case the estimator is highly variable, depending strongly on the data sample (see Figure 1 right). As noted in [1], the bias–variance trade off problem can be avoided by adopting the scale–space ideas which naturally lead to make inference on the smoothed curve  $f(\cdot; \nu)$  rather than on the curve  $f$ . It should be noted that, for small values of  $\nu$ , the smoothed curve  $f(\cdot; \nu)$  can be very different from  $f$ . However if  $\nu$  is within a reasonable range,  $f(\cdot; \nu)$ , which can be thought as the curve at a resolution level  $\nu$ , shows the same valley–peaks structure as  $f$  (see Figure 1, center). Thus, in order to assess the significance of features such as peaks and valleys, instead of constructing confidence intervals for  $f'(\theta)$ , CircSiZer seeks confidence intervals for the scale–space version  $f'(\theta; \nu)$ . Bootstrap– $t$  confidence intervals are constructed as detailed in [4].

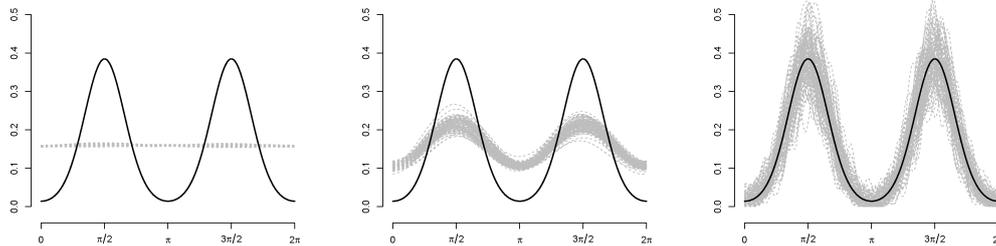


Figure 1: Mixture of two von Mises distributions,  $\nu M(\pi/2, 4)$  and  $\nu M(3\pi/4, 4)$ , in the same proportion (solid line) and nonparametric density estimators for the mixture from 100 random samples of size 250 (gray curves) with  $\nu = 0,2$  (left),  $\nu = 2$  (center) and  $\nu = 100$  (right).

The information provided by the construction of the intervals is displayed in a circular color map which reflects the statistical significance of the slope at  $(\theta, \nu)$  locations. So, for a given pair  $(\theta, \nu)$ , if the confidence interval is above (below) 0 then, the curve at location  $\theta$  and a smoothing level  $\nu$  is significantly increasing (decreasing), so that map location is colored blue (red). If the confidence interval contains 0, the curve at the smoothing level  $\nu$  and at the point  $\theta$  does not have a statistically significant slope and it is coloured in purple. Regions where there is not enough data to make statements about significance are gray coloured.

## Data analysis

In this section, the performance of CircSiZer map is illustrated with some simulated data sets and one real data set from an experiment with cemented femoral components.

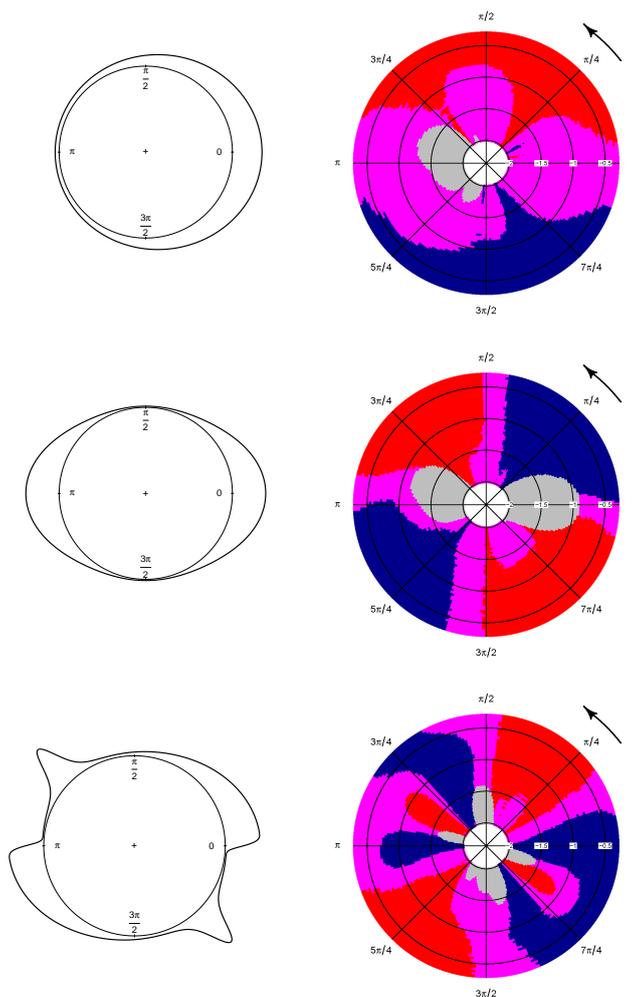


Figure 2: CircSiZer maps (right column) for kernel density estimates based on simulated data from a von Mises distribution (top-left), a mixture of two von Mises distributions (center-left) and a mixture of two wrapped skew-normal and two wrapped Cauchy distributions (bottom-left). Sample size  $n = 200$ . For reading CircSiZer, take as sense of rotation the direction marked by the arrow. Values of the smoothing parameter  $\nu$ , which are transformed to  $-\log(\nu)$  scale, are indicated along the radius.

As simulated models, the following distributions have been considered: a von Mises distribution, a mixture of two von Mises distributions and a mixture of two wrapped Cauchy and two wrapped skew-normal distributions. In Figure 2, the models (left column) and the corresponding CircSiZer maps (right column) are represented. The CircSiZer maps were obtained from a random sample of size  $n = 200$  from each model and a significance level of  $\alpha = 0,05$  was considered. Taking counterclockwise as the positive sense of rotation, Figure 2 (top-right) displays a blue area followed by a red area for a wide range of smoothing parameters, indicating a significant increase then decrease, i.e., unimodality. In Figure 2 (center-right), the bimodal structure is clearly brought out by the CircSiZer map, as the two peaks and the trough can be identified by the counterclockwise blue-red-blue-red pattern on the map that occurs for the whole range of smoothing parameters. In Figure 2 (bottom-right), it can be seen that only two modes are identified for values of the smoothing parameter smaller than  $\nu = 10$  but, for larger values of this parameter the cuatrimodal structure is obvious.

Now, the performance of CircSiZer is illustrated by the analysis of a real data set concerning angular positions of cracks in the cement mantle in a hip implant. The data set (described in more detail in [3]) is obtained from an in vitro fatigue study of six cemented total hip replacements using cadaveric femurs. Each specimen is loaded using a stair climbing apparatus and after loading, it is sectioned in 10 mm intervals from the level of the implant collar to the distal tip of the stem. For each section, angular positions of the cracks relative to the center of the stem section were documented. A counterclockwise definition of angular position of crack was used with a zero angle representing the lateral direction as shown in Figure 3 (left).

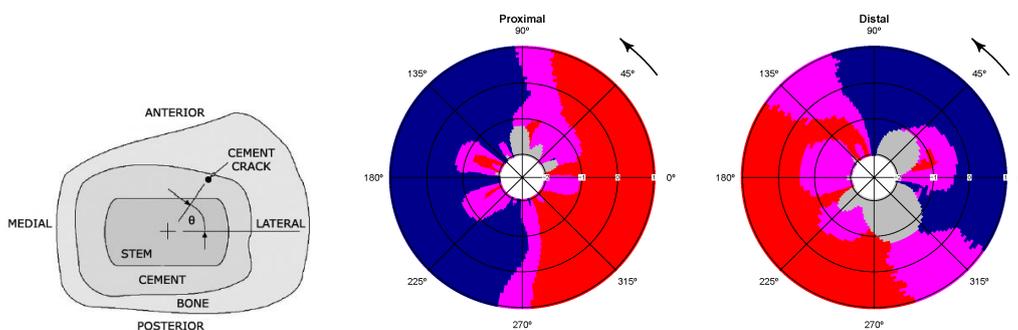


Figure 3: Left: A counterclockwise definition of angular position of crack was used with a zero angle representing the lateral direction. Center and right: CircSiZer maps for angular positions of cracks for one cemented implant for proximal (center) and distal (right) regions. Values of the smoothing parameter  $\nu$ , which are transformed to  $-\log_{10}$  scale, are indicated along the radius.

The goal of the study is to determine if cracks in the cement mantle are uniformly distributed or if there exists a directional dependence. The crack distribution for the proximal (sections at 10–50 mm) and distal regions (sections at 80–110 mm) of

the cement mantle is analyzed separately. The corresponding CircSiZer maps for the angular position of the cracks are shown in Figure 3 for proximal (center) and distal (right) regions which shows that the crack distribution around the cement mantle was not uniform in both cases. Moreover, the preferred direction is different, the cracks are concentrated around the posterior direction ( $270^\circ$ ) for proximal regions and around the anterior–medial direction of the mantle for distal regions.

## References

- [1] Chaudhuri, P. and Marron, J. S. (1999). *SiZer for exploration of structures in curves*, Journal of the American Statistical Association, **94**, pp. 807–823.
- [2] Di Marzio, M., Panzera A. and Taylor, C. C. (2009). *Local polynomial regression for circular predictors*, Statistics & Probability Letters, **79**, pp. 2066–2075.
- [3] Mann K. A., Gupta S., Race A., Miller, M. A., Cleary, R. J. (2003). *Application of circular statistics in the study of crack istribution around cemented femoral components*, Journal of Biomechanics **36**, pp. 1231–1234.
- [4] Oliveira, M., Crujeiras, R. M. and Rodríguez–Casal, A. (2013). *CircSiZer: an exploratory tool for circular data*, Journal of Environmental and Ecological Statistics, to appear.

## Solitones de Yamabe en grupos de Lie Lorentzianos de dimensión tres

Javier Seoane Bascoy

Departamento de Geometría y Topología

5 de junio de 2013

### Introducción

El objetivo de las ecuaciones de evolución geométrica es mejorar una métrica dada mediante un flujo asociado al objeto geométrico que estamos considerando. No obstante, existen métricas que, en vez de evolucionar con él, permanecen invariantes. Este es el caso de los solitones asociados a soluciones auto-similares del flujo.

El flujo de Yamabe, al igual que el flujo de Ricci, fue introducido por R. Hamilton. Dicho flujo deforma una variedad dada evolucionando su métrica mediante la ecuación  $\frac{\partial}{\partial t}g(t) = -\tau(t)g(t)$ , donde  $\tau(t)$  representa la curvatura escalar de la métrica  $g(t)$ . Una variedad pseudo-Riemanniana  $(M, g)$  es un *solitón de Yamabe* si y sólo si admite un campo de vectores  $X$  tal que

$$\mathcal{L}_X g = (\tau - \lambda)g,$$

donde  $\mathcal{L}_X$  denota la derivada de Lie en la dirección del campo de vectores  $X$  y  $\lambda$  es un número real.

En esta exposición estudiaremos la existencia de dichos solitones en los grupos de Lie Lorentzianos de dimensión tres.

### Variedades pseudo-Riemannianas

Antes de comenzar a estudiar el flujo de Yamabe y sus solitones asociados daremos algunos conceptos previos necesarios para nuestro trabajo. Consideramos  $(M, g)$  una *variedad pseudo-Riemanniana*  $n$ -dimensional, es decir, una variedad diferenciable de dimensión  $n$  dotada de un tensor métrico  $g$  (i.e. un tensor de tipo  $(0,2)$  simétrico y no degenerado) de signatura  $(\nu, n - \nu)$ . En particular, se dice que una variedad pseudo-Riemanniana es *Riemanniana* si el tensor métrico es definido positivo y es *Lorentziana* si su signatura es  $(1, n - 1)$ . Denotaremos por  $T_p M$  el espacio tangente a  $M$  en el punto  $p \in M$  y por  $\mathfrak{X}(M)$  el espacio de todos los campos de vectores diferenciables tangentes a  $M$ . Siguiendo la notación habitual en geometría

pseudo-Riemanniana, un vector distinto de cero  $z \in T_p M$  diremos que es *temporal* si  $g(z, z) < 0$ , *espacial* si  $g(z, z) > 0$  y *nulo* o *luminoso* si  $g(z, z) = 0$ .

En general podemos dotar a las variedades pseudo-Riemannianas de diferentes conexiones. Una *conexión* en una variedad  $M$  es un operador  $\nabla : \mathfrak{X}(M) \times \mathfrak{X}(M) \rightarrow \mathfrak{X}(M)$  verificando las siguientes propiedades:

- i)  $\nabla_X(fY_1 + gY_2) = X(f)Y_1 + f\nabla_X Y_1 + X(g)Y_2 + g\nabla_X Y_2$ ,
- ii)  $\nabla_{fX_1 + gX_2} Y = f\nabla_{X_1} Y + g\nabla_{X_2} Y$ .

Para cualesquiera  $X, X_1, X_2, Y, Y_1, Y_2 \in \mathfrak{X}(M)$  y  $f, g \in \mathfrak{F}(M)$ .

No obstante, entre todas las conexiones existe una de especial interés y que está asociada de forma directa al tensor métrico  $g$ , la llamada *conexión de Levi-Civita*. Dicha conexión, que denotaremos por  $\nabla$ , es la única conexión libre de torsión que paraleliza la métrica  $g$ ; i.e., dados  $X, Y \in \mathfrak{X}(M)$  la conexión  $\nabla$  verifica:

$$\mathcal{T}(X, Y) := \nabla_X Y - \nabla_Y X - [X, Y] = 0 \quad \text{y} \quad \nabla g = 0.$$

Además, su expresión explícita viene dada por la *fórmula de Koszul*:

$$\begin{aligned} 2g(\nabla_X Y, Z) &= X(g(Y, Z)) + Y(g(X, Z)) - Z(g(X, Y)) \\ &\quad + g(X, [Z, Y]) + g(Y, [Z, X]) + g(Z, [X, Y]), \end{aligned}$$

donde  $X, Y, Z \in \mathfrak{X}(M)$  y  $[\cdot, \cdot]$  representa el corchete de Lie.

Cada conexión tiene asociado de forma natural un tensor curvatura de tipo  $(1, 3)$ . En particular, el *tensor curvatura de tipo  $(1, 3)$*  asociado a la conexión de Levi-Civita esta dado por

$$R(X, Y)Z = \nabla_{[X, Y]} Z - \nabla_X \nabla_Y Z + \nabla_Y \nabla_X Z,$$

y su tensor curvatura de tipo  $(0, 4)$  asociado por  $R(X, Y, Z, V) = g(R(X, Y)Z, V)$ .

A partir de estos tensores curvatura y tomando una base arbitraria  $\{e_1, \dots, e_n\}$  de  $T_p(M)$  con  $g_{ij} = g(e_i, e_j)$ , se definen el *tensor de Ricci* y la *curvatura escalar* como

$$\rho(x, y) = \sum_{i,j=1}^n g^{ij} R(x, e_i, y, e_j), \quad \tau = \sum_{i,j=1}^n g^{ij} \rho(e_i, e_j), \quad (1)$$

respectivamente, donde  $(g^{ij})$  denota la matriz inversa de la métrica.

Como vimos, cada conexión determina una forma de definir la derivada direccional de un tensor; no obstante, existen otras formas de definir derivadas direccionales. Una de ellas es la *derivada de Lie*. La derivada de Lie de una función es la derivada direccional de la función, es decir, si  $f$  es una función real sobre  $M$  tenemos que  $\mathcal{L}_X f = X(f) = \nabla_X f$ . La derivada de Lie de un campo de vectores es el corchete de Lie, es decir, si  $Y$  es un campo de vectores  $\mathcal{L}_X Y = [X, Y]$ .

Extendemos la definición de derivada de Lie a un campo de tensores  $T$  de la siguiente forma:

$$(\mathcal{L}_X T)_p = \left. \frac{d}{dt} \right|_{t=0} ((\psi(t))^* T)_p,$$

donde  $\psi : I \times M \subset \mathbb{R} \times M \rightarrow M$  es el flujo local inducido por  $X$  y  $(\psi(t))^*$  es el pullback a lo largo del difeomorfismo  $\psi(t)$  para todo  $t \in I$ .

## Flujo de Yamabe

El flujo de Yamabe fue introducido por R. Hamilton al mismo tiempo que el flujo de Ricci. Dicho flujo deforma la variedad dada evolucionando su métrica mediante la ecuación

$$\frac{\partial}{\partial t}g(t) = -\tau(t)g(t), \quad (2)$$

donde  $\tau(t)$  representa la curvatura escalar de la métrica  $g(t)$ , a una nueva métrica con curvatura escalar constante dentro de la misma clase conforme.

Una familia 1-paramétrica de métricas  $g(t)$  definida a partir de una métrica inicial mediante homotecias y difeomorfismos, es decir, tal que

$$g(t) = \sigma(t)\psi(t)^*g(0),$$

donde  $\sigma(t)$  es una función positiva derivable y  $\psi(t) : M \rightarrow M$  es una familia 1-paramétrica de difeomorfismos verificando la ecuación (2) se llama *solución auto-similar*.

Si tomamos  $g(t)$  una solución auto-similar del flujo de Yamabe y la derivamos con respecto a  $t$  obtenemos la igualdad

$$-\tau(t)g(t) = \sigma'(t)\psi(t)^*g_0 + \sigma(t)\psi(t)^*(\mathcal{L}_X g_0), \quad (3)$$

donde  $g_0 = g(0)$ ,  $X$  es el campo de vectores dependiente del tiempo tal que  $X(\psi(t)(p)) = \frac{d}{dt}(\psi(t)(p))$  para cualquier  $p \in M$ , y  $\sigma' = \frac{d\sigma}{dt}$ .

Puesto que  $\tau(t)g(t) = \psi(t)^*\tau(0)g(0)$ , podemos suprimir los pullbacks en la ecuación (3) con lo que obtenemos

$$-\tau(0)g_0 = \sigma'(t)g_0 + \mathcal{L}_{\tilde{X}(t)}g_0,$$

donde  $\tilde{X}(t) = \sigma(t)X(t)$ . Evaluando en  $t = 0$  y tomando  $\lambda = -\sigma'(0)$  e  $Y_0 = -\tilde{X}(0)$  resulta que la ecuación se reduce a

$$\tau(0)g_0 = \lambda g_0 + \mathcal{L}_{Y_0}g_0.$$

Esto demuestra que para cualquier solución auto-similar del flujo de Yamabe existe un campo de vectores en  $M$  verificando

$$\mathcal{L}_Y g = (\tau - \lambda)g. \quad (4)$$

Recíprocamente, sea  $X$  un campo de vectores completo definido en una variedad pseudo-Riemanniana y denotemos por  $\psi(t) : M \rightarrow M$  con  $\psi(0) = id_M$  la familia de difeomorfismos generada por  $X$  verificando la relación

$$\frac{\partial}{\partial t}\psi(t)(p) = \frac{-1}{1 - \lambda t}X(\psi(t)(p)),$$

la cual está definida para todo  $t < \frac{1}{\lambda}$  si  $\lambda > 0$  y para todo  $t > \frac{1}{\lambda}$  si  $\lambda < 0$ . Considerando ahora la familia 1-paramétrica de métricas

$$g(t) = (1 - \lambda t)\psi(t)^*g,$$

tenemos que:

$$\begin{aligned} \frac{\partial}{\partial t}\{g(t)\} &= -\lambda\psi(t)^*g + (1 - \lambda t)\psi(t)^*(\mathcal{L}_{\frac{-1}{(1-\lambda t)}X}g) \\ &= \psi(t)^*(-\lambda g - \mathcal{L}_{X(\psi(t)(p))}g). \end{aligned}$$

Ahora, si el vector  $X$  verifica la ecuación (4), entonces

$$\frac{\partial}{\partial t}g(t) = -\psi(t)^*(\tau g) = -\tau(t)g(t),$$

lo que demuestra que  $g(t)$  es una solución del flujo de Yamabe dado por (2).

Una variedad pseudo-Riemanniana  $(M, g)$  admitiendo un campo de vectores  $X$  tal que

$$\mathcal{L}_X g = (\tau - \lambda)g,$$

donde  $\lambda$  es un número real, se denomina *solitón de Yamabe*. Además, dicho campo de vectores  $X$  verificando (4) se llama *campo de vectores solitón (para  $(M, g)$ )*. Un solitón de Yamabe se dice *contráctil*, *estable* o *expansivo* si admite un campo de vectores solitón para el cual  $\lambda > 0$ ,  $\lambda = 0$  ó  $\lambda < 0$ , respectivamente. Cuando  $(M, g)$  es, además, homogénea, puesto que en dichas variedades siempre existen campos de vectores Killing, diremos que un solitón de Yamabe es *no trivial* si admite un campo de vectores solitón que no es Killing. En particular, si la curvatura escalar no es constante entonces el solitón de Yamabe es necesariamente no trivial.

En el Teorema 4.1 de [3] Tashiro demuestra que una variedad Riemanniana completa de dimensión  $n \geq 2$  admite un campo de vectores homotético no Killing si y sólo si es localmente Euclídea; por tanto, en este caso, cualquier solitón de Yamabe no trivial con curvatura escalar constante es localmente Euclídeo. En lo que resta del trabajo demostraremos que el caso Lorentziano es mucho más rico. Concretamente, encontraremos grupos de Lie Lorentzianos de dimensión tres no llanos en los que existen ejemplos de dichos solitones.

### Solitones de Yamabe invariantes a la izquierda

Todo grupo de Lie  $(G, g)$  tiene asociado de forma natural un álgebra de Lie y viceversa. Para clasificar las álgebras de Lie de dimensión tres equipadas con un producto interior Lorentziano invariante haremos uso del producto vectorial inducido por los para-quaternios en el espacio  $\mathbb{R}_1^3$ , es decir, de la aplicación bilineal antisimétrica dada por las relaciones:

$$e_1 \times e_2 = -e_3, \quad e_3 \times e_1 = e_2, \quad e_2 \times e_3 = e_1;$$

donde  $\{e_1, e_2, e_3\}$  denota una base ortonormal de signatura  $(+ + -)$ . Con estas consideraciones se verifica el siguiente resultado de [2]:

**Teorema 1.** *El producto corchete en  $\mathfrak{g}$  está relacionado con el producto vectorial inducido por los para-quaternios mediante la fórmula  $[X, Y] = L(X \times Y)$ , donde  $L$  es una aplicación lineal de  $\mathfrak{g}$  en si misma definida de forma única. Además,  $\mathfrak{g}$  es unimodular si y solo si  $L$  es autoadjunta.*

Para el caso unimodular, puesto que en el caso Lorentziano un operador autoadjunto no es necesariamente diagonalizable debemos considerar sus posibles formas de Jordan asociadas. De este modo obtenemos las siguientes cuatro clases de álgebras de Lie unimodulares de dimensión tres:

Tipo Ia. Si  $L$  es diagonalizable con autovalores  $\{\alpha, \beta, \gamma\}$  con respecto a una base ortonormal  $\{e_1, e_2, e_3\}$  de signatura  $(++-)$ , entonces el correspondiente álgebra de Lie está dada por

$$(\mathfrak{g}_{Ia}) : \quad [e_1, e_2] = -\gamma e_3, \quad [e_1, e_3] = -\beta e_2, \quad [e_2, e_3] = \alpha e_1.$$

Tipo Ib. Supongamos que  $L$  tiene un autovalor complejo. Entonces, con respecto a una base ortonormal  $\{e_1, e_2, e_3\}$  de signatura  $(++-)$ , tenemos

$$L = \begin{pmatrix} \alpha & 0 & 0 \\ 0 & \gamma & -\beta \\ 0 & \beta & \gamma \end{pmatrix}, \quad \beta \neq 0,$$

y el álgebra de Lie correspondiente está dada por

$$(\mathfrak{g}_{Ib}) : \quad [e_1, e_2] = \beta e_2 - \gamma e_3, \quad [e_1, e_3] = -\gamma e_2 - \beta e_3, \quad [e_2, e_3] = \alpha e_1.$$

Tipo II. Supongamos que el polinomio minimal de  $L$  tiene una raíz de multiplicidad 2. Entonces, con respecto a una base  $\{e_1, e_2, e_3\}$  de signatura  $(++-)$ , tenemos

$$L = \begin{pmatrix} \alpha & 0 & 0 \\ 0 & \frac{1}{2} + \beta & -\frac{1}{2} \\ 0 & \frac{1}{2} & -\frac{1}{2} + \beta \end{pmatrix},$$

y el álgebra de Lie correspondiente está dada por

$$(\mathfrak{g}_{II}) : \quad [e_1, e_2] = \frac{1}{2}e_2 - (\beta - \frac{1}{2})e_3, \quad [e_1, e_3] = -(\beta + \frac{1}{2})e_2 - \frac{1}{2}e_3, \quad [e_2, e_3] = \alpha e_1.$$

Tipo III. Supongamos que el polinomio minimal de  $L$  tiene una raíz de multiplicidad 3. Entonces, con respecto a una base ortonormal  $\{e_1, e_2, e_3\}$  de signatura  $(++-)$ , tenemos

$$L = \begin{pmatrix} \alpha & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \alpha & 0 \\ -\frac{1}{\sqrt{2}} & 0 & \alpha \end{pmatrix},$$

y el álgebra de Lie correspondiente está dada por

$$(\mathfrak{g}_{III}) : \begin{cases} [e_1, e_2] = -\frac{1}{\sqrt{2}}e_1 - \alpha e_3, & [e_1, e_3] = -\frac{1}{\sqrt{2}}e_1 - \alpha e_2, \\ [e_2, e_3] = \alpha e_1 + \frac{1}{\sqrt{2}}e_2 - \frac{1}{\sqrt{2}}e_3. \end{cases}$$

Por otra parte, para las álgebras de Lie de dimensión tres no unimodulares y con curvatura seccional no constante existen las siguientes tres posibilidades:

$$(\mathfrak{g}_{IV}) : \quad [e_1, e_2] = 0, \quad [e_1, e_3] = \alpha e_1 + \beta e_2, \quad [e_2, e_3] = \gamma e_1 + \delta e_2;$$

donde  $\alpha + \delta \neq 0$  y, además, una de las siguientes condiciones se verifica:

IV.1  $\{e_1, e_2, e_3\}$  es ortonormal con  $\langle e_1, e_1 \rangle = -\langle e_2, e_2 \rangle = -\langle e_3, e_3 \rangle = -1$  y las constantes de estructura verifican  $\alpha\gamma - \beta\delta = 0$ .

IV.2  $\{e_1, e_2, e_3\}$  es ortonormal con  $\langle e_1, e_1 \rangle = \langle e_2, e_2 \rangle = -\langle e_3, e_3 \rangle = 1$  y las constantes de estructura verifican  $\alpha\gamma + \beta\delta = 0$ .

IV.3  $\{e_1, e_2, e_3\}$  es pseudo-ortonormal con

$$\langle \cdot, \cdot \rangle = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & -1 & 0 \end{pmatrix}$$

y las constantes de estructura verifican  $\alpha\gamma = 0$ .

Finalmente, estamos ya en condiciones de enunciar el teorema objetivo de este trabajo:

**Teorema 2.** *Un grupo de Lie Lorentziano no llano de dimensión tres es un solitón de Yamabe no trivial invariante a la izquierda si y solo si es localmente isométrico a un grupo de Lie no unimodular de Tipo IV.3 con  $\gamma = 0$  y  $\alpha = \frac{\delta}{2} \neq 0$ .*

## Referencias

- [1] Calviño-Louzao, E., Seoane-Bascoy, J., Vázquez-Abal, M. E. y Vázquez-Lorenzo, R. (2012). *Three-dimensional homogeneous Lorentzian Yamabe solitons*, Abhandlungen aus dem Mathematischen Seminar der Universität Hamburg, **82**, pp. 193–203.
- [2] Cordero, L. A. y Parker, Ph. (1997). *Left-invariant Lorentzian metrics on 3-dimensional Lie groups*, Rendiconti di Matematica e delle sue Applicazioni. Serie VII, **17**, pp. 129–155.
- [3] Tashiro, Y. (1965). *Complete Riemannian manifolds and some vector fields*, Transactions of the American Mathematical Society, **117**, pp. 251–275.

## La suma de todos los naturales y otras series divergentes

F. Adrián Fernández Tojo

Departamento de Análisis Matemático

19 de junio de 2013

### Introducción Histórica

#### Series convergentes

El uso de series se remonta por lo menos al 250–200 AEC. Por ejemplo, Arquímedes demostró que  $\sum_{n=0}^{\infty} \frac{1}{4^n} = \frac{4}{3}$  cuarteando el cuadrado. A pesar de que el concepto de **convergencia** no se establece formalmente hasta 1821 (Cauchy – *Cours d'Analyse Algébriques*), la idea de convergencia ya era familiar a Euler, Newton o Leibnitz. Sin embargo, los cálculos de estos matemáticos incluían a menudo series divergentes en ciertos pasos. Parece ser que Gauss fue el primero en mostrar un interés en las series divergentes más allá de su utilidad o aparición en los cálculos, pero no progresó en el campo.

#### Los inicios de las series divergentes

**Ejemplo 1** (Ejemplo de cálculo clásico).  $\sum_{n=0}^{\infty} z^n = \frac{1}{1-z}$  para  $z \in \mathbb{C}$ ,  $|z| < 1$ . La fórmula  $\frac{1}{1-z}$  está bien definida para todo  $z \in \mathbb{C}$ , salvo si  $z = 1$ . Por tanto  $\sum_{n=0}^{\infty} (-1)^n = \frac{1}{1-(-1)} = \frac{1}{2}$ .

*Sin embargo existen formas alternativas de definir esta serie:*

- $\sum_{n=0}^{\infty} (-1)^n = (1 - 1) + (1 - 1) + (1 - 1) \dots = 0$ ,
- $\sum_{n=0}^{\infty} (-1)^n = 1 + (-1 + 1) + (-1 + 1) + (-1 + 1) \dots = 1$ .

D’Alambert fue uno de los primeros en plasmar sus sospechas sobre estos cálculos en su carta de 1768.

Hardy, en [1], hace referencia a la situación de estos matemáticos que se enfrentaban a los comienzos de la teoría de series divergentes con las siguientes palabras:

Es cierto en general que los matemáticos anteriores a Cauchy no se preguntaban “¿cómo deberíamos definir  $1 - 1 + 1 - \dots$ ?”, sino “¿cuánto es  $1 - 1 + 1 - \dots$ ?” y que este hábito les llevó a confusiones y controversias que a menudo eran de un tono muy agresivo.

Hardy, *Series divergentes*, 1949.

### Principios filosóficos tras la definición de suma

A la hora de plantear una suma para el espacio de sucesiones sobre un grupo conmutativo  $G$ , esto es,  $\mathcal{S} = G^{\mathbb{N}}$ , surgen de forma natural las siguientes cuestiones:

- La suma, tal y como se entiende habitualmente, es la operación de  $G$ . Ésta define de manera natural una función  $S_n : G^n \rightarrow G$  de forma que  $S_n(x_1, \dots, x_n) = x_1 + \dots + x_n$ . Por lo tanto buscamos una función  $S : \mathcal{S} \rightarrow G$ .
- Además, cuando  $G$  es grupo con una una relación de orden (como en el caso de los números reales), se suele pedir que el grupo sea ordenado, esto es, que

$$x \leq y \Rightarrow x + z \leq y + z, \quad \forall x, y, z \in G,$$

o, equivalentemente, considerando la desigualdad elemento a elemento en  $G^n$ ,

$$x \leq y \Rightarrow S_n(x) \leq S_n(y), \quad \forall x, y \in G^n.$$

¿Podemos dotar a  $\mathcal{S}$  de una operación  $S$  con una propiedad análoga? Es decir, tal que:

$$x \leq y \Rightarrow S(x) \leq S(y), \quad \forall x, y \in \mathcal{S}.$$

Vamos a intentar definir, siguiendo estos principios, una operación  $S : \mathbb{R}_+^{\mathbb{N}} \rightarrow [0, \infty]$ , donde  $\mathbb{R}_+^{\mathbb{N}}$  es el espacio de sucesiones de números reales positivos, que satisfaga:

- *Orden*:  $x \leq y \Rightarrow S(x) \leq S(y), \quad \forall x, y \in \mathbb{R}_+^{\mathbb{N}}$ ,
- *Compatibilidad*:  $S(x_1, \dots, x_n, 0, \dots) = S_n(x_1, \dots, x_n)$ ,

maximizando además el conjunto sobre el cual la suma es finita.

Sea  $x = (x_1, x_2, \dots) \in \mathbb{R}_+^{\mathbb{N}}$ ,  $x^n = (x_1, \dots, x_n, 0, \dots)$ . La sucesión  $(S(x^n))_{n \in \mathbb{N}}$  es, por la propiedad de orden, monótona en  $[0, \infty]$ , luego tiene un límite  $s$ . Así,

$$S(x^n) \leq S(x) \leq +\infty, \quad \forall n \in \mathbb{N} \Rightarrow s \leq S(x) \leq +\infty.$$

Por lo tanto, resulta razonable tomar  $S(x) = s$ .

Esta suma (la suma usual de series) se puede extender de forma natural a series de números reales y complejos.

Sin embargo, la propiedad de orden puede ser demasiado exigente en general, pues no muchas series admiten una suma *usual*. Además, tenemos el siguiente teorema.

**Teorema 2** (Teorema de las series de Riemann). *Sea  $\sum_{n=1}^{\infty} a_n$  una serie convergente de números reales tal que  $\sum_{n=1}^{\infty} |a_n| = +\infty$ . Entoces, para todo  $s \in [-\infty, +\infty]$ , existe una permutación  $\sigma : \mathbb{N} \rightarrow \mathbb{N}$  tal que  $\sum_{n=1}^{\infty} a_{\sigma(n)} = s$ .*

El comienzo formal de la teoría de series divergentes surge con Cesàro, quien, en 1890, dio la primera definición generalizada de la suma de una serie.

**Definición 3.** Sea  $\sum_{n \in \mathbb{N}} a_n$  una serie infinita. Sea  $(s_n)_{n \in \mathbb{N}}$  su sucesión de sumas parciales y sea

$$t_n := \frac{1}{n} \sum_{k=1}^n s_k.$$

Entonces, si  $t_n \rightarrow s$  existe, entonces se dice que  $\sum_{n \in \mathbb{N}} a_n$  es sumable según Cesàro y  $\sum_{n \in \mathbb{N}} a_n = s$ .

La suma de Cesàro tiene la ventaja de ser intuitiva por ser la *suma esperada* de la serie.

En la época se sabía que el **producto de Cauchy** de dos series convergentes no tenía que ser convergente.

$$\left( \sum_{n \in \mathbb{N}} a_n \right) \left( \sum_{n \in \mathbb{N}} b_n \right) = \sum_{n \in \mathbb{N}} c_n, \text{ donde } c_n = \sum_{k=1}^n a_k b_{n-k}.$$

Sin embargo Cesàro probó lo siguiente.

**Teorema 4.** El producto de Cauchy de dos series convergentes es sumable según Cesàro.

## Métodos de suma

### Propiedades de los metodos

- **Regularidad:** Si el método de suma coincide con el usual en las series convergentes.
- **Linealidad:** Si  $r$  y  $s$  son sucesiones,  $k$  un escalar y  $S$  es un método de suma, entonces  $S(r + ks) = S(r) + kS(s)$  siempre que todas las sumas involucradas sean calculables por  $S$ .
- **Estabilidad:** Si  $s$  es una sucesión y  $s'_n = s_{n+1}$  entonces  $S(s) = s_0 + S(s')$ .
- **Re-indexabilidad finita:** Si cambiamos de lugar un número finito de elementos de una serie sumable por un método, el método proporciona la misma suma.

Se dice que dos metodos de suma  $A$  y  $B$  son **consistentes** si coinciden sobre las sucesiones sobre las cuales estan definidos.

En tal caso se dice que  $A$  es **más fuerte** que  $B$  si  $A$  está definido en toda sucesión sobre la que está definido  $B$ .

Calculemos  $\sum_{k=0}^{\infty} cr^k$  para  $c, r \in \mathbb{C}$ ,  $r \neq 1$ , usando algunos de los axiomas de estabilidad y linealidad:

$$G(r, c) := \sum_{k=0}^{\infty} cr^k = c + \sum_{k=0}^{\infty} cr^{k+1} = c + r \sum_{k=0}^{\infty} cr^k = c + rG(r, c),$$

por lo tanto

$$G(r, c) = \frac{c}{1-r}.$$

### Métodos por sucesiones

Los **métodos por sucesiones** consisten en asignar a la serie una sucesión cuyo límite, en caso de existir, se llamará suma de la serie.

Entre ellos están las sumas (o promedios) de Cesàro. Consideremos una serie  $a = \sum_{j=0}^{\infty} a_j$ . Definamos

$$A_n^{-1} = a_n, \quad A_n^{\alpha} = \sum_{k=0}^n A_k^{\alpha-1},$$

y sea  $E_n^{\alpha}$ ,  $A_n^{\alpha}$  para la serie  $1 + 0 + 0 + 0 \dots$ . Entonces la **suma de Cesàro generalizada**  $C_{\alpha}$  de  $a$  es

$$C_{\alpha}(a) := \lim_{n \rightarrow \infty} \frac{A_n^{\alpha}}{E_n^{\alpha}}.$$

Sea  $p = (p_n)_{n \in \mathbb{N}}$  una sucesión de números positivos tal que

$$\frac{p_n}{p_1 + \dots + p_n} \rightarrow 0.$$

Si ahora transformamos una sucesión  $a = (a_n)_{n \in \mathbb{N}}$  con los pesos  $(p_n)_{n \in \mathbb{N}}$  de la forma

$$t_m = \frac{p_m s_1 + p_{m-1} s_2 + \dots + p_1 s_m}{p_1 + \dots + p_m},$$

entonces, si  $t_m \rightarrow L \in \mathbb{R}$ ,  $L = N_p(a)$  se llama **promedio de Nördlund**  $N_p$  de  $a$ . Los promedios de Nördlund tienen las siguientes propiedades:

- Los promedios de Nördlund son regulares, lineales, estables y consistentes entre ellos.
- Si definimos  $p_n^k = \binom{n+k-1}{k-1}$ , entonces  $C_{\alpha}(a) = N_{(p^{\alpha})}(a)$ .
- Si  $h > k$ , entonces  $N_h$  es más fuerte que  $N_k$ .

## Métodos por funciones

Los **métodos por funciones** consisten en construir una función auxiliar (analítica o meromorfa en general) que nos permita asignar un valor a la serie.

El **método de Euler** consiste en asignar a una serie  $\sum_{n=1}^{\infty} a_n$  (posiblemente divergente) la función compleja  $f(x) := \sum_{n=1}^{\infty} a_n x^n$  que, en caso de ser analítica en un entorno de cero, podemos extenderla por continuación analítica. Si 1 pertenece al dominio de la extensión, decimos que **la suma de Euler** es  $\mathfrak{E}(\sum_{n=1}^{\infty} a_n) = f(1)$ .

La continuación analítica de una función no tiene por qué ser única, por tanto podríamos tener varios valores posibles para la suma.

### Ejemplo 5.

$$1 - 1 + 1 - 1 + \dots = \sum_{n=0}^{\infty} x^n \Big|_{x=-1} = \frac{1}{1-x} \Big|_{x=-1} = \frac{1}{2}.$$

### La función $\zeta$ de Riemann

Euler fue el primero en hacer uso de esta función para el cálculo de series divergentes (un siglo antes que Riemann).

La función  $\zeta$  se define como  $\zeta(s) := \sum_{n=1}^{\infty} n^{-s}$  y además admite extensión analítica en  $\mathbb{C} \setminus \{1\}$ . Por otra parte,

$$\zeta(-n) = -\frac{B_{n+1}}{n+1},$$

para  $n \in \mathbb{N}$  donde los  $B_n$  son los números de Bernoulli:

$$B_m := \sum_{k=0}^m \sum_{v=0}^k (-1)^v \binom{k}{v} \frac{v^m}{k+1}.$$

### Ejemplo 6.

$$1 + 2 + 3 + \dots = \sum_{n=1}^{\infty} \frac{1}{n^{-1}} = \zeta(-1) = -\frac{B_2}{2} = -\frac{1}{12}.$$

## Aplicaciones

### El efecto Casimir

El **efecto Casimir** o la **fuerza de Casimir–Polder** consiste en que dados dos objetos metálicos en el vacío, separados por una distancia pequeña comparada con el tamaño de los objetos, aparece una fuerza atractiva entre ambos.

En la interpretación de la **Teoría Cuántica de Campos** (QFT) el responsable de este efecto es el *vacío*. El vacío en QFT tiene las características de una partícula, pero, como es de esperar, casi todas estas características se cancelan en promedio.

Sin embargo, el valor esperado de la energía del vacío no es nulo.

Si consideramos el campo electromagnético cuantizado como un conjunto de infinitos osciladores armónicos simples cuya oscilación crea las ondas electromagnéticas, en su estado fundamental estos osciladores poseen algo de energía debido al principio de incertidumbre de Heisenberg.

Como cada oscilador sólo se corresponde con una frecuencia, tenemos infinitos osciladores en cada punto del espacio. Sumando la energía media de dichos osciladores obtenemos una cantidad infinita de energía.

Al colocar unas placas metálicas planas paralelas en el espacio, éstas limitan la cantidad de longitudes de onda que caben entre ellas, creando una diferencia de energía entre el exterior y el interior de las placas. Entre ambas placas sigue habiendo una cantidad infinita de energía pero, aún así, este infinito es *inferior* al infinito exterior.

Sean  $\hbar$  la constante reducida de Planck,  $c$  la velocidad de la luz,  $a$  la distancia entre las placas y  $A$  el área de las placas.

El cómputo de la diferencia entre las energías esperadas del vacío por unidad de área dentro y fuera de las placas es

$$\frac{\langle E(s) \rangle}{A} = \lim_{s \rightarrow 0} -\frac{\hbar c^{1-s} \pi^{2-s}}{2a^{3-s}} \frac{1}{3-s} \sum_n |n|^{3-s} = -\frac{\hbar c \pi^2}{6a^3} \zeta(-3) = \frac{-\hbar c \pi^2}{3 \cdot 240a^3}.$$

Así, la fuerza de Casimir por unidad de área que ejerce el vacío exterior sobre las placas es

$$\frac{F_c}{A} = -\frac{d}{da} \frac{\langle E \rangle}{A} = -\frac{\hbar c \pi^2}{240a^4}.$$

La fuerza es negativa, indicando pues el carácter atractivo de la misma: disminuyendo la distancia entre placas, la energía es reducida.

## Referencias

- [1] Hardy, G. H. (1949). *Divergent series*, Oxford.
- [2] Shawyer, B. y Watson B. (1994). *Borel's Methods of Summability. Theory and Applications*, Oxford.
- [3] Hawking, S. (2001). *El universo en una cáscara de nuez*, Crítica.
- [4] Casimir, H. B. G. (1948). *On the attraction between two perfectly conducting plates*, **51**, pp. 793–795.



## ¿Por qué giran los electrones?

Santiago Codesido Sánchez

Departamento de física de partículas, USC

26 de junio de 2013

### Introducción

#### Historia del spin

En 1922 el experimento de Stern–Gerlach cumplió uno de los objetivos de la física de partículas de principio de siglo, medir el giro de los electrones en los átomos. Estos son cargas en rotación que por tanto se comportan como un pequeño imán, al que medimos la dirección.

Se esperaba una distribución uniforme, como parecería indicar la intuición clásica. El resultado, sin embargo, resultó ser que el giro toma un conjunto *discreto* de valores. Es decir, no hay electrones *quietos*.

La explicación de este hecho se encuentra en los aspectos geométricos de la teoría de campos con la que podemos modelar al electrón.

### Campos

#### Formulación matemática

**Definición 1.** Sea una variedad  $T$  a la que llamamos *base*, y un espacio vectorial  $C$  de valores. Definimos un *campo* como una aplicación  $\mathbf{q} \in \mathcal{C}^\infty(T, C)$ .

Consideraremos por simplicidad que  $T$  es un espacio vectorial, ya que todo el cálculo que sigue se puede generalizar con facilidad a variedades a través del fibrado tangente<sup>2</sup>.

Ejemplos comunes de campo son:  $T = \mathbb{R}$ ,  $C = \mathbb{R}^3$ , la posición de una partícula en el tiempo;  $T = \mathbb{R}^3$ ,  $C = \mathbb{R}^3$ , las velocidades de un fluido; o  $T = \mathbb{R}^{3,1}$ ,  $C = \mathbb{C}$ , la función de onda de Schrödinger de una partícula.

Buscamos una noción de *velocidad* generalizando la idea de derivada de la posición. Queremos una derivada, por tanto, de campos.

---

PALABRAS CLAVE: spin; teoría de campos; representaciones; álgebras de Lie.

<sup>2</sup>Esta generalización pasa por dotar al fibrado de una *conexión*, lo cual se escapa de los objetivos del presente texto. Se puede hallar una descripción mucho más general en [1].

**Definición 2.** Sea  $\mathbf{u} : T \rightarrow T$ , definimos la derivada direccional respecto de  $\mathbf{u}$  como

$$\nabla_{\mathbf{u}} : \mathcal{C}^{\infty}(T, C) \rightarrow \mathcal{C}^{\infty}(T, C)$$

$$\mathbf{q}(x) \mapsto \lim_{h \rightarrow 0} \frac{\mathbf{q}(x + h\mathbf{u}(x)) - \mathbf{q}(x)}{h}.$$

Sean ahora  $f : T \rightarrow \mathbb{R}$ ,  $\mathbf{u}, \mathbf{v} : T \rightarrow T$  y  $\mathbf{q} : \mathcal{C}^{\infty}(T, C)$ . Se cumple entonces el siguiente resultado:

**Proposición 3.** La derivada direccional es lineal respecto al campo que se toma,  $(\nabla_{\mathbf{u}+f\mathbf{v}}\mathbf{q})(x) = (\nabla_{\mathbf{u}}\mathbf{q})(x) + f(x) (\nabla_{\mathbf{v}}\mathbf{q})(x)$ .

**Definición 4.** Definimos una **derivada de campos**  $\nabla : \mathcal{C}^{\infty}(T, C) \rightarrow \mathcal{C}^{\infty}(T, C \otimes T^*)$  dada por la condición  $(\nabla\mathbf{q})(x)[\mathbf{u}(x)] = (\nabla_{\mathbf{u}}\mathbf{q})(x)$  con  $\mathbf{q} \in \mathcal{C}^{\infty}(T, C)$  y  $\mathbf{u} : T \rightarrow T$  en virtud de la linealidad de  $\nabla_{\mathbf{u}}$  en  $\mathbf{u}$ .

**Definición 5.** Definimos la **velocidad de un campo**  $\mathbf{q}$  como  $\mathbf{v} = \nabla\mathbf{q} \in \mathcal{C}^{\infty}(T, C \otimes T^*)$ .

## Representaciones de la rotación<sup>3</sup>

### Relaciones de conmutación

Las rotaciones vienen dadas por un eje  $\mathbf{n} = (n_x, n_y, n_z)$ ,  $\|\mathbf{n}\| = 1$  y un ángulo  $\theta \in [-\pi, \pi]$ , y las notamos por  $R(\mathbf{n}, \theta) \in \text{SO}(3)$ . Es decir, las rotaciones son elementos del grupo especial ortogonal de 3 dimensiones.

Tal y como las hemos definido, una propiedad evidente de las rotaciones es que  $R(\mathbf{n}, \pi) = R(\mathbf{n}, -\pi)$ , que como veremos más adelante es la que determina las peculiaridades de su topología.

Una serie de ejemplos de campos que son representación de  $\text{SO}(3)$  son los campos de temperatura (representación trivial o escalar), los campos de velocidades (representación fiel o vectorial) y los campos de tensiones (representación tensorial).

Su álgebra viene dada por  $\mathfrak{so}(3) = \langle \{\sigma_x, \sigma_y, \sigma_z\} \rangle$  (donde  $x, y, z$  hacen referencia a los tres ejes ortogonales de  $\mathbb{R}^3$ ), de modo que  $\sigma_{\mathbf{n}} = \sum_i n_i \sigma_i$  genera a  $R(\mathbf{n}, \theta) = \exp\{\theta\sigma_{\mathbf{n}}\}$  con la condición  $\exp\{2\pi\sigma_i\} = \text{Id}$ . La operación del álgebra de Lie está dada por las siguientes relaciones de conmutación  $[\sigma_x, \sigma_y] = \sigma_z$ ,  $[\sigma_z, \sigma_x] = \sigma_y$ ,  $[\sigma_y, \sigma_z] = \sigma_x$ .

Tal y como la hemos construido, como variedad  $\text{SO}(3)$  es una bola de dimensión 3 y radio  $\pi$  con los puntos opuestos identificados, es decir,  $\text{SO}(3) \equiv \mathbb{RP}^3$ . Conocemos así su grupo fundamental, que es  $\pi_1(\text{SO}(3)) = \mathbb{Z}_2$ . Por otra parte,  $\mathfrak{so}(3)$  es isomorfa al álgebra  $\mathfrak{su}(2)$  del grupo de Lie  $\text{SU}(2)$ , que como variedad es  $\mathbb{S}^3$ , lo cual se puede ver a través de la identificación con los cuaterniones unitarios. El hecho de que  $\text{SU}(2)$  sea simplemente conexo, estando generado por la misma álgebra que  $\text{SO}(3)$ , nos lleva a que es el recubrimiento universal de  $\text{SO}(3)$ , también llamado  $\text{Spin}(3)$ .

<sup>3</sup>Para más información sobre grupos y álgebras de Lie, así como de sus representaciones, que se usarán en las siguientes secciones, se puede consultar [2].

## Representación espinorial

$\mathfrak{su}(2)$  tiene, salvo isomorfismo, la siguiente representación fiel (llamadas matrices de Pauli), actuando sobre el espacio vectorial  $\mathbb{C}^2$ :

$$\sigma_x = \frac{i}{2} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \sigma_y = \frac{i}{2} \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \sigma_z = \frac{i}{2} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

El factor  $\frac{1}{2}$  es fundamental para mantener las relaciones de conmutación anteriormente indicadas. Esto implica que todas tienen dos autovectores  $S^\pm$  con autovalores  $\pm \frac{i}{2}$ , al igual que, por linealidad, tiene todo  $\sigma_{\mathbf{n}}$  arbitrario. A su vez, implica que  $\exp\{2\pi\sigma_i\} = -\text{Id}$ , lo que se sigue fácilmente de la fórmula de Euler. Es decir, la transformación que corresponde a un giro completo en este espacio *no es la identidad*, sino la reflexión.

Como veremos luego, son los campos en esta representación los que causan el spin del electrón. Por ello, los llamamos *espinoriales*.

## Representación sobre funciones

Consideremos coordenadas cilíndricas en  $\mathbb{R}^3$ :  $(r, z, \theta)$  respecto a un eje  $\mathbf{n}$  dado por la coordenada  $z$ .

Buscamos estudiar las rotaciones respecto al eje  $\mathbf{n}$ , de tal modo que nos interesa la dependencia de las funciones en el ángulo. Sea  $f \in C^\infty(\mathbb{R}/(2\pi\mathbb{Z}), \mathbb{C})$  una función compleja en el ángulo  $\theta$ .

Queremos que  $\sigma_{\mathbf{n}}$  actúe en el espacio vectorial de funciones como  $\exp\{\tilde{\theta}\sigma_{\mathbf{n}}\} \cdot f(\theta) = f(\theta + \tilde{\theta})$ . Es decir, trasladando el ángulo.

Por el teorema de Taylor y la expansión exponencial<sup>4</sup> tenemos que  $f(\theta) + \tilde{\theta}(\sigma_{\mathbf{n}}f)(\theta) = f(\theta + \tilde{\theta}) + R(\tilde{\theta})$  con  $\lim_{\theta \rightarrow 0} \frac{R(\theta)}{\theta} = 0$ . Vemos entonces que la representación de los giros sobre las funciones es  $\sigma_{\mathbf{n}} = \partial_\theta$ .

## Dinámica de campos

### Formalismo Lagrangiano

#### Lagrangiano y acción

**Definición 6.** Llamamos **Lagrangiano** a una aplicación  $L : C \times (C \otimes T^*) \rightarrow \mathbb{R}$ .

**Definición 7.** Llamamos **acción de Hamilton** al funcional sobre los campos  $\mathbf{q}$  dado por

$$S[\mathbf{q}] = \int_T L(q(x), \nabla q(x)) dx.$$

<sup>4</sup>Para grupos de Lie que admiten un álgebra unitaria que genera al corchete de Lie mediante el conmutador.

**Postulado.** *El Lagrangiano codifica la dinámica de un determinado sistema, de tal modo que las ecuaciones del movimiento para un campo  $\mathbf{q}$  vienen dadas por los extremos de la acción.*

### Fuerzas y momentos

**Definición 8.** Sean  $V, W$  espacios vectoriales,  $F : V \rightarrow W$ , y  $u, v, w \in V$ . Definimos la **derivada de Gâteaux** como el límite, si existe,  $\left. \frac{\partial F(u)}{\partial u} \right|_{u=v} [w] := \lim_{\epsilon \rightarrow 0} \frac{F(v+\epsilon w) - F(v)}{\epsilon}$ .

**Proposición 9.** Debido a la linealidad en  $w$ , tenemos que  $\left. \frac{\partial F(u)}{\partial u} \right|_{u=v} \in \mathcal{L}(V, W)$ .

**Definición 10.** Sea  $\mathbf{q} : C^\infty(T, C)$  y  $L : C \times (C \otimes T^*) \rightarrow \mathbb{R}$ . Definimos el momento de  $\mathbf{q}$ ,  $\mathbf{p}_{\mathbf{q}}$ , y la fuerza de  $\mathbf{q}$ ,  $\mathbf{F}_{\mathbf{q}}$  como

$$\mathbf{p}_{\mathbf{q}} : T \rightarrow \mathcal{L}(C \otimes T^*, \mathbb{R}) = \mathcal{L}(C, T) \quad \mathbf{F}_{\mathbf{q}} : T \rightarrow \mathcal{L}(C, \mathbb{R})$$

$$x \mapsto \left. \frac{\partial L(\tilde{q}, \tilde{v})}{\partial \tilde{v}} \right|_{(\tilde{q}, \tilde{v})=(\mathbf{q}(x), \nabla \mathbf{q}(x))}, \quad x \mapsto \left. \frac{\partial L(\tilde{q}, \tilde{v})}{\partial \tilde{q}} \right|_{(\tilde{q}, \tilde{v})=(\mathbf{q}(x), \nabla \mathbf{q}(x))}.$$

### Ecuaciones del movimiento

**Teorema 11** (Euler–Lagrange). Si  $q$  tiene soporte compacto y  $\mathbf{F}_{\mathbf{q}}$  y  $\mathbf{p}_{\mathbf{q}}$  están bien definidas como derivadas, el extremo de la acción  $S$  se corresponde a  $\nabla \cdot \mathbf{p}_{\mathbf{q}} = \mathbf{F}_{\mathbf{q}}$ .

Esta ecuación es una generalización de la segunda ley de Newton.

### Simetrías y corrientes conservadas

**Definición 12.** Sea  $s : T \rightarrow \mathfrak{gl}(C)$ ,  $\alpha \in \mathbb{R}$  y  $S_\alpha : T \rightarrow \text{GL}(C)$  de modo que  $S_\alpha(x) = \exp\{\alpha s(x)\}$ . Decimos que  $s$  es una **simetría continua** de  $S[\mathbf{q}]$  si  $\exists a \in \mathbb{R}^+$  tal que  $S[\mathbf{q}] = S[S_\alpha \cdot \mathbf{q}]$ ,  $\forall \alpha \in [-a, a]$ .

**Definición 13.** Definimos la **corriente** de  $\mathbf{q}$  por  $s$ ,  $\mathbf{j}_{\mathbf{q}_s} : T \rightarrow T$ , como  $\mathbf{j}_{\mathbf{q}_s}(x) = \mathbf{p}_{\mathbf{q}}(x)[s(x) \cdot \mathbf{q}(x)]$ .

**Teorema 14** (Noether). Si estamos en condiciones de aplicar el teorema de Euler–Lagrange y  $s$  es una simetría continua, entonces  $\nabla \cdot \mathbf{j}_{\mathbf{q}_s} = 0$  y decimos que  $\mathbf{j}_{\mathbf{q}_s}$  es una **corriente conservada**.

Es interesante en muchas ocasiones considerar cantidades conservadas integrales, en el sentido de aislar una sola dimensión del espacio base  $T$  integrando el resto. Sea así  $T = \mathbb{R}_t \oplus T'$  de modo que  $\nabla = \partial_t \oplus \nabla|_{T'}$  y  $\mathbf{j}_{\mathbf{q}_s}(t) \in L^2(T', C)$ .

Integramos  $\nabla \cdot \mathbf{j}_{\mathbf{q}_s} = 0^5$ , usando el teorema de Stokes, entonces

$$-\partial_t \int_{T'} \mathbf{j}_{\mathbf{q}_s}|_{\mathbb{R}_t} dx = \int_{T'} \nabla|_{T'} \cdot \mathbf{j}_{\mathbf{q}_s}|_{T'} dx = \int_{T'} d\alpha = \int_{\partial T'} \alpha = 0,$$

donde  $\alpha$  es una  $(\dim(T') - 1)$ -forma y  $\partial T' = \emptyset$ .

<sup>5</sup>Usamos que podemos conmutar  $\int_{T'}$  con  $\partial_t$ .

**Definición 15.** Definimos la cantidad integral  $\mathbf{J}_{\mathbf{q}_s}(t) = \int_{T'} \mathbf{j}_{\mathbf{q}_s}|_{\mathbb{R}_t}(t, x) dx \in \mathcal{C}^\infty(\mathbb{R}_t, \mathbb{R})$  de modo que si  $\mathbf{j}_{\mathbf{q}_s}$  es una corriente conservada,  $\partial_t \mathbf{J}_{\mathbf{q}_s} = 0$  y decimos que es una *cantidad integral conservada*.

## Momento angular

### Partícula puntual

Para relacionar el momento angular, o *cantidad de giro* con las cantidades conservadas, consideremos primero un caso sencillo y conocido. Sea un campo que describe las coordenadas planas  $\mathbf{q} : \mathbb{R} \rightarrow \mathbb{R}^2$ ,  $q(t) = (r(t), s(t))$  de modo que  $r$  es la distancia al centro,  $s$  es la distancia sobre la circunferencia, contando también el número de vueltas y  $j = mv_s r = p_s r$  es el momento angular o cantidad de giro.

Veamos un ejemplo de como calcular una cantidad conservada. Sea  $\sigma \in \mathfrak{gl}(\mathbb{R}^2)$  el generador de los giros. Ha de cumplir  $\exp\{\phi\sigma\} \cdot (r, s) = (r, s + r\phi)$ . Entonces

$$\sigma = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix},$$

de lo que se sigue  $\sigma^2 = 0$  y  $\exp\{\phi\sigma\} = \text{Id} + \phi\sigma$ . Un giro arbitrario es entonces  $\exp\{\phi\sigma\} \cdot (r, s) = (r, s) + \phi(0, r) = (r, s + r\phi)$  como era de esperar.

Asumiendo que los giros son una simetría del sistema, calculamos el momento angular con el teorema de Noether,  $\mathbf{j}_{\mathbf{q}_\sigma} = \mathbf{p}_{\mathbf{q}}[\sigma \cdot \mathbf{q}]$ . El momento lineal  $\mathbf{p}_{\mathbf{q}}$  lo podemos expresar por  $\mathbb{R}^{2*} = \mathbb{R}^2 \implies \mathbf{p}_{\mathbf{q}} = (p_r, p_s)$ . Entonces vemos que la cantidad conservada asociada a los giros es en efecto el mismo momento angular que se define en los cursos de física básica:

$$\mathbf{j}_{\mathbf{q}_\sigma} = \begin{pmatrix} p_r & p_s \end{pmatrix} \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} r \\ s \end{pmatrix} = p_s r = L.$$

## Campo espinorial

### Representación de los giros

Sea un campo en coordenadas cilíndricas  $(t, r, z, \theta)$  respecto a un eje  $\mathbf{n}$ ,  $\mathbf{q} \in \mathcal{C}^\infty(\mathbb{R}_t \times \mathbb{R}_r^+ \times \mathbb{R}_z \times \mathbb{R}_\theta / (2\pi\mathbb{Z}), \mathbb{C}) \otimes \mathbb{C}^2$ . Podemos descomponerlo en autofunciones de  $\sigma_{\mathbf{n}}$ ,  $\mathbf{q}_{m\pm} = c(t, r, z) e^{im\theta} S_\pm$ , donde  $S_\pm$  son los autovectores de la representación de  $\text{SU}(2)$  y  $e^{im\theta}$  con  $m \in \mathbb{Z}$  las autofunciones de la representación  $\partial_\theta$ .

Entonces

$$\begin{aligned} \sigma_{\mathbf{n}} \mathbf{q}_{m\pm} &= c(t, r, z) \sigma_{\mathbf{n}} \left( e^{im\theta} S_\pm \right) \\ &= c(t, r, z) \left( \sigma_{\mathbf{n}} e^{im\theta} \right) S_\pm + c(t, r, z) e^{im\theta} (\sigma_{\mathbf{n}} S_\pm) = i \left( m \pm \frac{1}{2} \right) \mathbf{q}_{m\pm}. \end{aligned}$$

Aplicamos el teorema de Noether  $l_{m\pm} = \mathbf{j}_{\mathbf{q}_{m\pm}\sigma_{\mathbf{n}}} = \mathbf{p}_{\mathbf{q}_{m\pm}}[\sigma_{\mathbf{n}} \cdot \mathbf{q}_{m\pm}]$ , y  $l_{m\pm} = [m \pm \frac{1}{2}] i \mathbf{p}_{\mathbf{q}_{m\pm}}[\mathbf{q}_{m\pm}]$ . Vemos que  $m$  se debe a la componente *espacial* del momento angular y  $\pm \frac{1}{2}$  a la componente *intrínseca*, es decir, la parte entera se debe a la

dependencia espacial mientras que la semientera se debe característicamente a la representación espinorial.

Para calcular la integral conservada necesitamos el lagrangiano de Dirac, que encierra la cinética de una onda que describe al electrón<sup>6</sup>:  $L(q, v) = -iq^* \cdot \hat{\sigma} \cdot v$ , donde  $\hat{\sigma} \in \text{GL}(\mathbb{C}^2) \otimes \mathbb{R}^4$  es el “vector de Dirac”, dado por  $(\hat{\sigma}, v_t \oplus \mathbf{v}) = v_t \text{Id}_{\mathbb{C}^2} + \|\mathbf{v}\| \sigma_{\hat{\mathbf{v}}}$ . El lagrangiano es invariante bajo rotaciones (se sigue de las relaciones de conmutación de las  $\sigma$ ). Entonces el momento es  $\mathbf{p}_{\mathbf{q}} = -i\mathbf{q}^* \hat{\sigma}$  y  $\mathbf{j}_{\mathbf{q}m \pm \sigma_n} = [m \pm \frac{1}{2}] \mathbf{q}^* \hat{\sigma} \mathbf{q}$ . Como la restricción de  $\hat{\sigma}$  a  $\mathbb{R}_t$  es la identidad, finalmente  $\mathbf{J}_{\mathbf{q}m \pm \sigma_n} = (m \pm \frac{1}{2}) \int_{\mathbb{R}^3} \mathbf{q}^* \cdot \mathbf{q} dx$ .

### Cuantización

En la mecánica cuántica hallamos una interpretación probabilística de la función de onda, de tal modo que la integral de la densidad de probabilidad ha de ser la unidad:  $\int_{\mathbb{R}^3} \mathbf{q}^* \cdot \mathbf{q} d^3x = 1$ . En este caso

$$\mathbf{J}_{\mathbf{q}m \pm \sigma_n} = \left( m \pm \frac{1}{2} \right).$$

Es notable recordar que *no hemos recurrido a la mecánica cuántica hasta ahora*. Es decir, la cuantización del spin *es* de hecho un fenómeno cuántico, pero el origen del spin es un fenómeno *geométrico*.

La medida en la mecánica cuántica requiere que el resultado sea un autoestado del observable. Midamos el eje que midamos siempre obtendremos una respuesta del tipo  $(m \pm \frac{1}{2})$ , lo que tal y como observó el experimento de Stern–Gerlach indica que *no hay electrones quietos*.

Por supuesto, hay partículas en otras representaciones de  $\text{SO}(3)$  dadas por su *spin* como  $\pi$  con spin 0,  $\gamma$  con spin 1 o incluso  $\Delta^{++}$  con spin 3/2. El hecho de que los campos de spin 1/2 sean la materia se debe al teorema spin–estadística, que dice que los campos de spin semientero obedecen al principio de exclusión de Pauli, que es lo que hace a la materia característicamente impenetrable, por contra de los campos de fuerzas, que pueden acumularse sin límite y están representados por campos de spin entero.

### Referencias

- [1] Nakahara, M. (2003). *Geometry, topology and physics*, IOP Publishing.
- [2] Baradarajam, V. S. (1984). *Lie groups, Lie algebras and their representations*, Springer.
- [3] Weinberg, S. (2000). *The quantum theory of fields*, Cambridge University Press.

<sup>6</sup>Estrictamente describe a un electrón *quiral*, el Lagrangiano completo describe a las dos quiralidades. Una descripción mucho más detallada se puede encontrar en [3].

## Ecuaciones diferenciales de orden fraccionario

Óscar A. Otero Zarrquiños

Departamento de Análisis Matemático

3 de julio de 2013

### Introducción

En este trabajo presentamos una introducción a las ecuaciones diferenciales de orden fraccionario y el análisis del movimiento de un proyectil en términos del cálculo fraccional.

### Preliminares

De acuerdo con la aproximación de Riemann–Liouville al cálculo fraccional, la noción de integral fraccionaria de orden  $\alpha$  ( $\alpha > 0$ ) es una consecuencia natural de la fórmula de Cauchy:

$$I^n f(t) := \frac{1}{(n-1)!} \int_0^t (t-s)^{n-1} f(s) ds, \quad t > 0, \quad n \in \mathbb{N}.$$

**Definición 1.** La *integral fraccionaria de orden*  $\alpha > 0$  de una función  $f \in L^1([a, b], \mathbb{R})$  se define como:

$$I_{a+}^\alpha f(t) = \int_a^t \frac{(t-s)^{\alpha-1}}{\Gamma(\alpha)} f(s) ds,$$

donde  $\Gamma$  es la función gamma Euler.

**Teorema 2.** Sean  $f(t)$  y  $g(t)$  tales que  $I^\alpha f(t)$ ,  $I^\alpha g(t)$ ,  $I^{\alpha+\beta} g(t)$  existen. Entonces, las siguientes propiedades básicas de la integral de Riemann–Liouville se cumplen:

1. Interpolación (continuidad).
2. Linealidad.
3. Propiedad de semigrupo (ley de los exponentes).
4. Conmutatividad.

---

PALABRAS CLAVE: Caputo; Riemann–Liouville; cálculo fraccional; proyectil.

Denotando por  $D^n$  con  $n \in \mathbb{N}$ , el operador de la derivada de orden  $n$ , primero notemos que

$$D^n I^n = I \quad \text{pero} \quad I^n D^n \neq I, \quad n \in \mathbb{N},$$

i.e., tiene inversa por la izquierda (y no inversa por la derecha) del operador integral correspondiente  $I^n$ .

**Definición 3.** Para una función  $f$  dada en un intervalo  $[a, b]$ , la **derivada fraccionaria de orden  $\alpha$  de Riemann–Liouville** de  $f$  se define como:

$$D^\alpha f(t) = \frac{1}{\Gamma(n - \alpha)} \left( \frac{d}{dt} \right)^n \int_a^t (t - s)^{-\alpha + n - 1} f(s) ds,$$

donde  $n = [\alpha] + 1$  y  $[\alpha]$  es la parte entera de  $\alpha$ .

Algunas propiedades:

$$D^\alpha I^\alpha = I, \quad \alpha \geq 0,$$

$$D^\alpha t^\gamma = \frac{\Gamma(\gamma + 1)}{\Gamma(\gamma + 1 - \alpha)} t^{\gamma - \alpha}, \quad \alpha > 0, \quad \gamma - 1, t > 0,$$

$$D^\alpha 1 = \frac{t^{-\alpha}}{\Gamma(1 - \alpha)}, \quad \alpha > 0, \quad t > 0.$$

**Definición 4.** Sea  $f \in AC^n([a, b])$ , siendo  $AC^n([a, b])$  el conjunto de funciones  $f$  tales que  $f^i$  es absolutamente continua  $\forall i = 0, \dots, n$ . La **derivada fraccionaria de Caputo** de  $f$  viene dada por:

$${}^c D^\alpha f(t) := \frac{1}{\Gamma(n - \alpha)} \int_a^t (t - s)^{n - \alpha - 1} f^{(n)}(s) ds.$$

De acuerdo con esta definición, una propiedad relevante es que la derivada fraccionaria de una constante sigue siendo cero, i.e.:

$${}^c D^\alpha 1 = 0, \quad \alpha > 0. \quad (1)$$

Exploramos ahora las principales diferencias entre ambas definiciones dadas para las derivadas fraccionarias. Para la derivada fraccionaria de Riemann–Liouville:

$$D^\alpha (t - a)^{\alpha - j} = 0, \quad j = 1, 2, \dots, [\alpha] + 1. \quad (2)$$

De las propiedades (1) y (2) reconocemos los siguientes enunciados sobre funciones que para  $t > 0$  admiten la misma derivada fraccionaria de orden  $\alpha$ , con  $n - 1 < \alpha < n, n \in \mathbb{N}$ , es decir:

$$D^\alpha f(t) = D^\alpha g(t) \Leftrightarrow f(t) = g(t) + \sum_{j=1}^n c_j (t - a)^{\alpha - j},$$

$${}^c D^\alpha f(t) = {}^c D^\alpha g(t) \Leftrightarrow f(t) = g(t) + \sum_{j=1}^n c_j (t - a)^{n - j}.$$

## Análisis del movimiento de un proyectil en términos del cálculo fraccional

### Definiciones y preliminares

La **transformada de Laplace de una función**  $f(t)$  definida para todos los números positivos  $t \geq 0$ , es la función  $F(s)$ , definida por:

$$F(s) = L\{f(t)\} = \int_0^{\infty} e^{-st} f(t) dt,$$

siempre que esté bien definido.

Si le aplicamos la transformada de Laplace a la derivada fraccionaria de Caputo de una función  $f$  nos da:

$$L \left\{ \frac{d^\alpha}{dx^\alpha} f(x) \right\} = s^\alpha F(s) - \sum_{m=0}^{n-1} s^{\alpha-m-1} f^{(m)}(0), \quad n-1 < \alpha < n. \quad (3)$$

Si le aplicamos la transformada de Laplace a la derivada fraccionaria de Riemann-Liouville de una función  $f$  nos da:

$$L \left\{ \frac{d^\alpha}{dx^\alpha} f(x) \right\} = s^\alpha F(s) - \sum_{m=0}^{n-1} s^m f^{(\alpha-m-1)}(0), \quad n-1 < \alpha < n. \quad (4)$$

### Formulación del problema

Las ecuaciones clásicas del movimiento de una partícula en el plano  $x-y$  vienen dadas por:

$$m \frac{d^2 x}{dt^2} = 0, \quad m \frac{d^2 y}{dt^2} = -mg, \quad (5)$$

según la clásica Ley de Newton, y con las condiciones iniciales:

$$\begin{aligned} x(0) &= 0, & \dot{x}(0) &= v_0 \cos \phi, \\ y(0) &= 0, & \dot{y}(0) &= v_0 \operatorname{sen} \phi. \end{aligned}$$

es decir, el proyectil parte del reposo, con una fuerza inicial de módulo  $v_0$  y un ángulo  $\phi$ .

Podemos escribir las ecuaciones diferenciales fraccionarias para el problema del proyectil de la siguiente forma:

$$\frac{d^\alpha x}{dt^\alpha} = 0, \quad \frac{d^\alpha y}{dt^\alpha} = -g, \quad \text{donde } 1 < \alpha \leq 2, \quad (6)$$

donde para  $\alpha = 2$  tenemos el caso clásico (5).

Para la **derivada fraccionaria de Caputo**, se tiene que la solución de (6) se puede obtener mediante la transformada de Laplace (3) como sigue:

$$\begin{aligned} s^\alpha X(s) - s^{\alpha-1}x(0) - s^{\alpha-2}\dot{x}(0) &= 0, \\ s^\alpha Y(s) - s^{\alpha-1}y(0) - s^{\alpha-2}\dot{y}(0) &= -\frac{g}{s}. \end{aligned} \quad (7)$$

Usando las condiciones iniciales (7), se llega a que:

$$X(s) = \frac{v_0 \cos \phi}{s^2}, \quad Y(s) = -\frac{g}{s^{\alpha+1}} + \frac{v_0 \operatorname{sen} \phi}{s^2}.$$

Por tanto:

$$x(t) = v_0 \cos \phi \cdot t, \quad y(t) = -\frac{gt^\alpha}{\Gamma(\alpha+1)} + v_0 \operatorname{sen} \phi \cdot t. \quad (8)$$

Para la **derivada fraccionaria de Riemann–Liouville**, se tiene que la solución de (6) puede ser hallada usando (4):

$$\begin{aligned} s^\alpha X(s) - D^{\alpha-1}x(0) - sD^{\alpha-2}x(0) &= 0, \\ s^\alpha Y(s) - D^{\alpha-1}y(0) - sD^{\alpha-2}y(0) &= -\frac{g}{s}. \end{aligned}$$

Usando ahora las condiciones iniciales:

$$\begin{aligned} D^{\alpha-2}x(0) &= 0, & D^{\alpha-1}x(0) &= v_0 \cos \phi, \\ D^{\alpha-2}y(0) &= 0, & D^{\alpha-1}y(0) &= v_0 \operatorname{sen} \phi. \end{aligned}$$

que para  $\alpha = 2$  coincide con las condiciones iniciales (7). Obtenemos:

$$X(s) = \frac{v_0 \cos \phi}{s^\alpha}, \quad Y(s) = -\frac{g}{s^{\alpha+1}} + \frac{v_0 \operatorname{sen} \phi}{s^\alpha}.$$

De ahí se sigue que:

$$x(t) = v_0 \cos \phi \frac{t^{\alpha-1}}{\Gamma(\alpha)}, \quad y(t) = -\frac{gt^\alpha}{\Gamma(\alpha+1)} + v_0 \operatorname{sen} \phi \frac{t^{\alpha-1}}{\Gamma(\alpha)}. \quad (9)$$

### Características del movimiento de un proyectil en el cálculo fraccional

Tres cantidades son particularmente relevantes para identificar, distinguir y analizar trayectorias: rango, altura máxima y tiempo de vuelo.

- El **rango** es la distancia horizontal recorrida por el proyectil desde que se lanza hasta que aterriza.
- La **altura máxima** es la altura en el punto más alto de la trayectoria.
- El **tiempo de vuelo** es la cantidad de tiempo que el proyectil está en el aire desde que se lanza hasta que aterriza.

## Trayectoria

### Caputo

Eliminando  $t$  de (8), obtenemos la **trayectoria del proyectil fraccional** en un  $\alpha$  arbitrario:

$$y = \tan \phi \cdot x - \frac{gx^\alpha}{\Gamma(\alpha + 1)(v_0 \cos \phi)^\alpha}. \quad (10)$$

Si  $\alpha \rightarrow 2$ , (10) nos da la ecuación clásica de la trayectoria:

$$y = \tan \phi \cdot x - \frac{g}{2v_0^2} \sec^2 \phi \cdot x^2.$$

### Riemann–Liouville

Eliminando ahora  $t$  de (9), se obtiene:

$$y = \tan \phi \cdot x - \frac{g}{\Gamma(\alpha + 1)} \left( \frac{x\Gamma(\alpha)}{v_0 \cos \phi} \right)^{\frac{\alpha}{\alpha-1}}. \quad (11)$$

Nuevamente si  $\alpha \rightarrow 2$ , a partir de (11) deducimos la ecuación clásica de la trayectoria:

$$y = \tan \phi \cdot x - \frac{g}{2v_0^2} \sec^2 \phi \cdot x^2.$$

## Rango

El **rango proyectil fraccional** se define como el valor de  $x$  en el punto de impacto.

### Caputo

Así,  $y = 0$  en  $x = R_F$ . De ahí,  $R_F$  viene dado por:

$$R_F = (v_0)^{\frac{\alpha}{\alpha-1}} \left[ \frac{\Gamma(\alpha + 1)}{g} \right]^{\frac{1}{\alpha-1}} (\sen \phi)^{\frac{1}{\alpha-1}} \cos \phi. \quad (12)$$

También si  $\alpha \rightarrow 2$ , (12) nos lleva al rango del proyectil clásico:

$$R_C = \frac{2v_0^2}{g} \sen \phi \cos \phi.$$

### Riemann–Liouville

Haciendo nuevamente  $y = 0$  en  $x = R_F$ , se obtiene:

$$R_F = \frac{v_0 \cos \phi}{\Gamma(\alpha)} \left( \frac{\alpha v_0 \sen \phi}{g} \right)^{\alpha-1}. \quad (13)$$

Si  $\alpha \rightarrow 2$ , a partir de (13) obtenemos el rango del proyectil clásico:

$$R_C = \frac{2v_0^2}{g} \sen \phi \cos \phi.$$

### Tiempo de vuelo

El **tiempo de vuelo fraccional**  $t_{F\text{-flight}}$  se define como el valor de  $t$  en el cual el proyectil choca con el suelo.

#### Caputo

Por esto hacemos,  $y = 0$  en  $t = t_{F\text{-flight}}$ , y de ahí:

$$t_{F\text{-flight}} = \left[ \frac{\Gamma(\alpha + 1)v_0 \text{sen } \phi}{g} \right]^{\frac{1}{\alpha-1}}. \quad (14)$$

#### Riemann–Liouville

De modo similar obtenemos:

$$t_{F\text{-flight}} = \frac{\alpha v_0 \text{sen } \phi}{g}. \quad (15)$$

Nótese que en ambos casos, el tiempo de vuelo clásico puede deducirse de las ecuaciones (14) y (15) cuando  $\alpha \rightarrow 2$ :

$$t_{C\text{-flight}} = \frac{2v_0 \text{sen } \phi}{g}.$$

### Altura máxima

#### Caputo

El proyectil alcanza su **altura máxima** cuando la componente vertical desaparece, i.e,  $\dot{y} = 0$ . Resolviendo esta ecuación para  $t$ , llegamos a que:

$$t_{\text{flight}-h} = \left[ \frac{\Gamma(\alpha)v_0 \text{sen } \phi}{g} \right]^{\frac{1}{\alpha-1}}. \quad (16)$$

Substituyendo (16) en  $y(t)$  en (8), obtenemos  $H_F$ :

$$H_F = \left( 1 - \frac{1}{\alpha} \right) \left[ \frac{\Gamma(\alpha)}{g} \right]^{\frac{1}{\alpha-1}} (v_0 \text{sen } \phi)^{\frac{\alpha}{\alpha-1}}.$$

#### Riemann–Liouville

De modo similar obtenemos:

$$t_{\text{flight}-h} = \frac{(\alpha - 1)v_0 \text{sen } \phi}{g},$$

$$H_F = \frac{-((\alpha - 1)v_0 \text{sen } \phi)^\alpha}{\Gamma(\alpha + 1)g^{\alpha-1}} + \frac{(v_0 \text{sen } \phi)^\alpha}{\Gamma(\alpha)} \left( \frac{\alpha - 1}{g} \right)^{\alpha-1}.$$

### Referencias

- [1] Abdelhalin, E. (2011). *Analysis of projectile motion view of fractional calculus*, Applied Mathematical Modelling, **35**, pp. 1231–1239.

## ¿Cómo contrastar la independencia estadística?

Eduardo García Portugués

Departamento de Estadística e Investigación Operativa

17 de julio de 2013

El concepto de dependencia entre variables aleatorias ha jugado siempre un papel central en la Estadística. Por ello, es imprescindible contar con tests estadísticos que indiquen cuándo dos variables son independientes o dependientes. Este trabajo se estructura en dos partes: en la primera se hace un breve repaso a los tests estadísticos más relevantes para contrastar la independencia y en la segunda se propone un nuevo test basado en la estimación no paramétrica de la densidad. Este nuevo test se aplica para estudiar los incendios de Portugal de los últimos 25 años.

### Motivación

Contrastar la independencia estadística es importante para determinar si existe o no una relación entre dos variables aleatorias de forma rigurosa, objetiva y automática. Es, además, un primer paso en el análisis estadístico de un conjunto de variables: si se concluye que hay independencia, el tratamiento de las variables puede ser marginal, mientras que si no se deberá explorar el tipo de dependencia (por ejemplo mediante modelos de regresión). Finalmente, es conveniente tener claro que la intuición puede jugar malas pasadas a la hora de examinar visualmente la independencia entre dos variables. Por ejemplo, la Figura 1 contiene los diagramas de dispersión de muestras obtenidas de diferentes variables  $X$  e  $Y$ . Situaciones que a priori parecen dependientes (4 y 5) resultan ser independientes y, recíprocamente, situaciones que parecen independientes resultan no serlo (3 y 6). Este hecho refuerza la necesidad de tener tests estadísticos rigurosos para distinguir estas situaciones.

### Tests de independencia

Un test estadístico es una función de la muestra  $X_1, \dots, X_n$  de una variable aleatoria  $X$  que nos va a permitir decidir si una hipótesis  $H_0$  sobre  $X$  es plausible a la luz de la muestra. Los conceptos clave de todo test estadístico son los siguientes:

- Hipótesis nula o de partida  $H_0$ : es cierta hasta que haya evidencias que demuestren lo contrario.

---

PALABRAS CLAVE: independencia; test estadístico; densidad.

- Estadístico  $T$ : mide la discrepancia entre  $H_0$  cierta y falsa.
- Distribución de  $T$  bajo  $H_0$ : permite saber si la discrepancia, suponiendo que  $H_0$  es cierta, es grande (rechazo  $H_0$ ) o pequeña (no rechazo  $H_0$ ).
- $p$ -valor: probabilidad de obtener un estadístico  $T$  al menos tan desfavorable a  $H_0$  como el observado, suponiendo que  $H_0$  es cierta. Se rechaza  $H_0$  si el  $p$ -valor es menor que el nivel de significación  $\alpha$ , usualmente  $\alpha = 0,05$ .

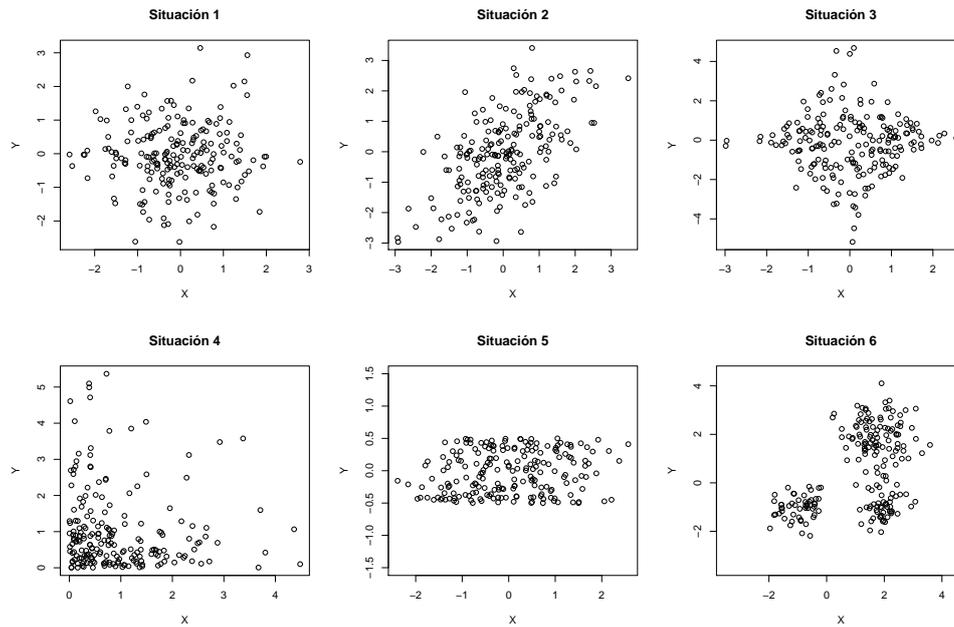


Figura 1: Diagramas de dispersión de 200 observaciones simuladas de  $X$  e  $Y$ . En las situaciones 1, 4 y 5 son independientes, mientras que en 2, 3 y 6 no.

Intuitivamente, que dos variables sean independientes significa que ninguna afecta en el comportamiento de la otra. La definición formal es la siguiente.

**Definición 1** (Independencia estadística). *Dos variables aleatorias  $X$  e  $Y$  son **independientes** si  $F_{XY}(x, y) = F_X(x)F_Y(y)$ , siendo  $F_{XY}$  la función de distribución conjunta de  $X$  e  $Y$  y  $F_X$  y  $F_Y$  las distribuciones marginales correspondientes. Si  $X$  e  $Y$  son absolutamente continuas, esto equivale a  $f_{XY}(x, y) = f_X(x)f_Y(y)$ , siendo  $f_{XY}$  la densidad conjunta y  $f_X$  y  $f_Y$  las densidades marginales.*

Sin suponer nada más sobre las variables  $X$  e  $Y$  (salvo que sean absolutamente continuas), en la literatura estadística han surgido cuatro grandes alternativas para contrastar la independencia:

1. Usar la **función de distribución**, considerando el proceso empírico dado por la diferencia de distribuciones empíricas,  $\sqrt{n}(F_n(x, y) - F_n(x)F_n(y))$ .

2. Test **Chi-cuadrado**. Consiste en discretizar los datos en categorías y realizar un test Chi-cuadrado en una tabla de contingencia.
3. Usar **medidas de asociación**. Son los métodos más populares por su interpretación y facilidad de uso. Se estiman mediante las medidas muestrales de asociación.
4. Usar la **función de densidad**, estimándola no paramétricamente mediante suavizado núcleo (kernel smoothing).

Nos centraremos en las dos últimas al ser las más relevantes para la siguiente sección.

### Tests basados en medidas de asociación

Las medidas de asociación miden el grado de dependencia de primer orden (o en media) existente entre variables. La medida de asociación más conocida es el *coeficiente de correlación de Pearson*, que corresponde al conocido coeficiente de correlación de la recta de regresión de  $Y$  sobre  $X$ , denotado por  $r = \text{corr}(X, Y)$ . Al estar determinado por una recta, este coeficiente suele ser bastante limitado y por ello han surgido modificaciones como la *Rho de Spearman*, que es el coeficiente de correlación de las probabilidades acumuladas:  $\rho = \text{corr}(F(X), G(Y))$ . Otra medida de asociación es la *Tau de Kendall*, que mide el grado de concordancia ( $X$  grande  $\implies Y$  grande) o discordancia ( $X$  grande  $\implies Y$  pequeña):  $\tau = 2(\mathbb{P}\{(X_1 - X_2)(Y_1 - Y_2) > 0\} - \mathbb{P}\{(X_1 - X_2)(Y_1 - Y_2) < 0\}) - 1$ .

Hay procedimientos para contrastar que cualquiera de estas medidas sea nula ( $H_0 : r, \rho, \tau = 0$ , no asociación). Sin embargo, conviene tomarlos con cautela a la hora de contrastar la independencia, ya que aunque ésta implica que  $H_0 : r, \rho, \tau = 0$ , el recíproco no es cierto. Entre otras situaciones, esto ocurre si la dependencia entre las variables es de segundo orden (o en varianza).

### Tests basados en la densidad

Rosenblatt [5] propone un estadístico utilizando la estimación por suavizado para las densidades desconocidas. El estadístico mide la discrepancia de la independencia mediante la distancia  $\mathcal{L}_2$  entre la estimación de la densidad conjunta y el producto de las estimaciones marginales:

$$R_n = \int_{\mathbb{R}^2} \left( \hat{f}_{XY;h}(x, y) - \hat{f}_{X;h}(x) \hat{f}_{Y;h}(y) \right)^2 dx dy,$$

donde  $\hat{f}_{X;h}(x) = \frac{1}{nh} \sum_{i=1}^n K\left(\frac{x-X_i}{h}\right)$  (respectivamente  $\hat{f}_{Y;h}(y)$ ) y  $\hat{f}_{XY;h}(x, y) = \frac{1}{nh^2} \sum_{i=1}^n K\left(\frac{x-X_i}{h}\right) \times K\left(\frac{y-Y_i}{h}\right)$  son los estimadores núcleo de las densidades desconocidas (ver [3] para referencias en este campo). Debido al diseño del estadístico  $R_n$ , el test detecta la presencia de dependencia de una forma plena, a diferencia de los test basados en asociación, aunque ni Rosenblatt [5] ni posteriores referencias proporcionan un procedimiento aplicable en la práctica o estudios de simulación.

## Un nuevo test de independencia

El test de Rosenblatt puede ser considerado para contrastar la independencia en otros contextos más complejos, como por ejemplo, en el que una variable sea escalar y la otra *direccional*. Los datos direccionales representan observaciones contenidas en la  $q$ -esfera  $\Omega_q = \{\mathbf{x} \in \mathbb{R}^{q+1} : \|\mathbf{x}\| = 1\}$ , siendo casos particulares los datos circulares ( $q = 1$ ) y los datos esféricos ( $q = 2$ ). Hay ejemplos de este tipo de datos en diversas áreas como proteómica, meteorología, astronomía y ciencias forestales. Debido a la geometría de  $\Omega_q$ , las técnicas estadísticas usuales no valen para datos direccionales, siendo el ejemplo típico que la media de  $358^\circ$  y  $2^\circ$  no es  $0^\circ$ .

Sean  $\mathbf{X}$  y  $Z$  dos variables aleatorias absolutamente continuas,  $\mathbf{X}$  con soporte en  $\Omega_q$  y  $Z$  en  $\mathbb{R}$ . La hipótesis de independencia se puede expresar mediante

$$H_0 : f_{\mathbf{X}Z}(\mathbf{x}, z) = f_{\mathbf{X}}(\mathbf{x})f_Z(z) \text{ vs. } H_1 : f_{\mathbf{X}Z}(\mathbf{x}, z) \neq f_{\mathbf{X}}(\mathbf{x})f_Z(z).$$

El estadístico de contraste, al estilo del propuesto por [5], es

$$T_n = \int_{\Omega_q \times \mathbb{R}} \left( \hat{f}_{\mathbf{X}Z;h,g}(\mathbf{x}, z) - \hat{f}_{\mathbf{X};h}(\mathbf{x})\hat{f}_{Z;g}(z) \right)^2 \omega_q(d\mathbf{x}) dz,$$

donde  $\hat{f}_{\mathbf{X}Z;h,g}(\mathbf{x}, z) = \frac{c_{h,q}(L)}{ng} \sum_{i=1}^n L\left(\frac{1-\mathbf{x}^T \mathbf{X}_i}{h^2}\right) \times K\left(\frac{z-Z_i}{g}\right)$  es el estimador núcleo direccional-lineal conjunto y  $\hat{f}_{\mathbf{X};h}(\mathbf{x}) = \frac{c_{h,q}(L)}{n} \sum_{i=1}^n L\left(\frac{1-\mathbf{x}^T \mathbf{X}_i}{h^2}\right)$  y  $\hat{f}_{Z;g}(z) = \frac{1}{ng} \sum_{i=1}^n K\left(\frac{z-Z_i}{g}\right)$  son los estimadores núcleo direccional y lineal marginales. La explicación detallada de estos estimadores se puede ver en [3] y en sus referencias.

Bajo ciertas condiciones de regularidad se puede obtener la distribución asintótica de  $T_n$  bajo  $H_0$ ,

$$(nh^q g)^{\frac{1}{2}} (T_n - A_n) \xrightarrow{d} \mathcal{N}(0, 2\sigma_I^2),$$

siendo  $A_n$  y  $\sigma_I^2$  el sesgo y la varianza asintótica, respectivamente.

### El test en la práctica

Desgraciadamente, la distribución asintótica no sirve para calibrar el estadístico en la práctica debido a su lenta convergencia. Es por ello que es necesario implementar un procedimiento de remuestreo para aproximar la distribución real del estadístico a partir de una muestra  $\{(\mathbf{X}_i, Z_i)\}_{i=1}^n$  de la variable  $(\mathbf{X}, Z)$ . La idea es sencilla: bajo independencia,  $T_n$  debería ser similar al estadístico calculado a partir de una muestra obtenida mediante una permutación  $\sigma \in S_n$ ,  $\{(\mathbf{X}_i, Z_{\sigma(i)})\}_{i=1}^n$ . Mediante este método podemos obtener  $B \leq n!$  estadísticos permutados con los que es posible aproximar la verdadera distribución de  $T_n$ . Estos pasos se detallan en el siguiente algoritmo.

**Algoritmo 2** (Procedimiento del test). Sea  $\{(\mathbf{X}_i, Z_i)\}_{i=1}^n$  una muestra de  $(\mathbf{X}, Z)$ .

1. Calcular el estadístico  $T_n$ .

2. *Calibración por permutaciones. Para  $b = 1, \dots, B$ :*

- *Para una permutación aleatoria  $\sigma$ , obtener  $\{(\mathbf{X}_i, Z_{\sigma(i)})\}_{i=1}^n$ .*
- *Calcular el estadístico para la muestra permutada  $T_n^{*b}$ .*

3. *Obtener  $p$ -valor  $\approx \#\{T_n \leq T_n^{*b}\}/B$ .*

El cálculo de  $T_n$  y su calibración se pueden acelerar notablemente utilizando la expresión cerrada,

$$T_n = \frac{1}{E_q(h)^2} \mathbf{1} \left( \frac{1}{n^2} \Psi(h) \circ \Omega(g) - \frac{2}{n^3} \Psi(h) \Omega(g) + \frac{1}{n^4} \Psi(h) \mathbf{1}^T \mathbf{1} \Omega(g) \right) \mathbf{1}^T,$$

donde  $\Psi(h)$  y  $\Omega(g)$  son matrices simétricas y  $E_q(h) = (2\pi)^{\frac{q+1}{2}} h^{q-1} \mathcal{I}_{\frac{q-1}{2}}(1/h^2)$ . La gran ventaja es que sólo es necesario calcular una vez las matrices aunque en el calibrado se usen  $B$  permutaciones. La expresión es válida para los núcleos  $L$  y  $K$  más habituales: von Mises y normal, respectivamente.

### Estudio de simulación

En [3] se detalla un completo estudio de simulación del nuevo test para varias dimensiones, tamaños muestrales y diferentes escenarios. Además se compara con los test existentes para el caso circular–lineal ( $q = 1$ ):  $R_n^2$  de Mardia [4],  $U_n$  de Mardia [4] y  $\lambda_{4n}$  de Fisher y Lee [2]. Estos test están basados en adaptaciones a datos circular–lineales de las medidas de asociación  $r$ ,  $\rho$  y  $\tau$ , respectivamente.

En el caso circular–lineal, los resultados del estudio de simulación indican que el nuevo test es competitivo frente a  $R_n^2$ ,  $U_n$  y  $\lambda_{4n}$  para los tipos de asociación para los cuales estos tests fueron diseñados. Además, el nuevo test es capaz de detectar tipos de dependencia tanto en primer como en segundo orden que ninguno de los tests competidores son capaces de detectar. En el caso esférico–lineal el test funciona también de forma satisfactoria y no tiene competidores.

### Aplicación a datos reales

En [1] se analizan los 26870 incendios ocurridos en Portugal durante los años 1985 y 2005 y se clasifican en 102 cuencas diferentes. La recolección de estos datos se realiza comparando las imágenes satelitales de Portugal al principio y al final de la temporada de incendios. Del perímetro de cada incendio se puede obtener la recta que maximiza la varianza de los datos proyectados en ella y que forma su esqueleto principal, denominada primera componente principal (CP1). La CP1 induce un ángulo, que es la orientación del incendio, como puede verse en el gráfico de la izquierda de la Figura 2. Por el mismo método se puede calcular la orientación de una cuenca. Varias preguntas surgen para estos datos: ¿influye la orientación del incendio en el área quemada?, ¿influye la orientación de una cuenca en el área quemada dentro de ella?, ¿cómo es la dependencia entre la orientación de incendios y área quemada por cuencas? El test propuesto contesta a estas preguntas.

La hipótesis nula de **independencia** entre la orientación de incendios y su (log) área quemada se **rechaza** con  $p$ -valor = 0,000. Se rechaza también la independencia entre la orientación de la cuenca y el (log) área total quemada en ella para orientaciones con  $p$ -valor = 0,000. En cuanto a la dependencia por cuencas, el mapa de la Figura 2 ilustra que la **dependencia no es homogénea** en las cuencas y que está concentrada en unas pocas. Los tests se han realizado con ventanas elegidas por validación cruzada y con  $B = 1000$  permutaciones.

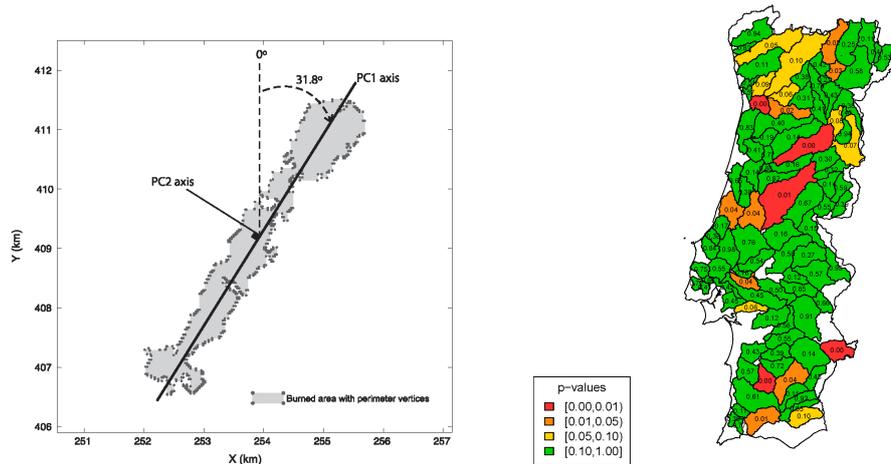


Figura 2: Izquierda: orientación angular del incendio por medio de la CP1 del perímetro (extraído de [1]). Derecha:  $p$ -valores del test de independencia con los incendios agregados por cuencas.

## Referencias

- [1] Barros, A. M. G., Pereira, J. M. C. y Lund, U. J. (2012). *Identifying geographical patterns of wildfire orientation: A watershed-based analysis*, Forest Ecology and Management, **264**, pp. 98–107.
- [2] Fisher, N. I. y Lee, A. J. (1981). *Nonparametric measures of angular-linear association*, Biometrika, **68**(3), pp. 629–636.
- [3] García-Portugués, E., Barros, A. M. G., Crujeiras, R. M., González-Manteiga, W. y Pereira, J. M. C. (2013). *A test for directional-linear independence, with applications to wildfire orientation and size*, Stochastic Environmental Research and Risk Assessment, to appear.
- [4] Mardia, K. V. (1976). *Linear-circular correlation coefficients and rhythmometry*, Biometrika, **63**(2), pp. 403–405.
- [5] Rosenblatt, M. (1975). *A quadratic measure of deviation of two-dimensional density estimates and a test of independence*, The Annals of Statistics, **3**(1), pp. 1–14.

## On an inverse spectral problem for the Sturm–Liouville operator in a finite interval

Nizomjon Jumaniyazov

Department of Applied Mathematics, USC  
 Department of Mathematical Physics and Applied Mathematics, UrSU

23th July, 2013

### Abstract

In this paper an inverse problem for the Sturm–Liouville operator in a finite interval is studied and some important uniqueness theorems are presented. We also focus on the construction of a Sturm–Liouville boundary value problem by studying two sequences satisfying exact conditions via the Gelfand–Levitan algorithm. We also show an exact example using this algorithm.

The following boundary value problem is called the Sturm–Liouville problem:

$$Ly \equiv -y'' + q(x)y = \lambda y, \quad x \in [0, \pi], \quad (1)$$

$$\begin{cases} y'(0) - hy(0) = 0, \\ y'(\pi) + Hy(\pi) = 0, \end{cases} \quad (2)$$

where  $q(x) \in C([0, \pi])$  is a real continuous function,  $h$  and  $H$  are given real numbers and  $\lambda$  is a complex parameter.

If (1) is considered with the boundary condition  $y(0) = 0, \quad y(\pi) = 0$ , then the resulting problem is called the **Dirichlet** problem. If (1) is considered with the boundary conditions  $y'(0) = 0, \quad y'(\pi) = 0$ , then the resulting problem is called the **Neumann** problem.

We denote by  $\varphi(x, \lambda)$  the solution of the equation in (1) satisfying the following initial conditions:

$$\varphi(0, \lambda) = 1, \quad \varphi'(0, \lambda) = h.$$

**Definition 3.** *If for some  $\lambda = \lambda_0$ , the problem (1)–(2) has a nontrivial solution  $y(x, \lambda_0)$ , then the number  $\lambda_0$  is called an **eigenvalue** of the problem (1)–(2) while the function  $y(x, \lambda_0)$  is called an **eigenfunction**.*

**Definition 4.** *The numbers*

$$\alpha_n^2 = \int_0^\pi \varphi^2(x, \lambda_n) dx, \quad n = 0, 1, 2, \dots,$$

---

KEYWORDS: Sturm–Liouville operator; Uniqueness theorems; Gelfand–Levitan algorithm.

are called *normalizing constants* of the problem (1)–(2).

**Definition 5.** The functions determined by the formulas

$$u_n(x) = \frac{1}{\alpha_n} \varphi(x, \lambda_n), \quad n = 0, 1, 2, \dots,$$

are called *orthonormal eigenfunctions* of the problem (1)–(2).

From now on, we call the pair of sequences  $\{\lambda_n\}_{n=0}^{\infty}$  and  $\{\alpha_n\}_{n=0}^{\infty}$  **spectral characteristics** or **spectral data** of the Sturm–Liouville problem.

**Definition 6.** The problem of finding spectral data  $\{\lambda_n\}_{n=0}^{\infty}$  and  $\{\alpha_n\}_{n=0}^{\infty}$  by the potential function and boundary conditions, is called the **direct spectral problem** and, reversely, the problem of constructing the boundary value problem (1)–(2), that is, to find  $q(x)$  and numbers  $h$  and  $H$ , is called the **inverse spectral problem**.

## Inverse problems

The first result which played an important role in the development of inverse problems for Sturm–Liouville operator was first obtained by V. A. Hambardzumyan in 1929 (see [1]).

We consider the so called Sturm–Liouville problem:

$$Ly \equiv -y'' + q(x)y = \lambda y, \quad y'(0) = 0, \quad y'(\pi) = 0. \quad (3)$$

where  $q(x) \in C([0, \pi])$  is a real continuous function. It is well known that if  $q(x) \equiv 0$ ,  $x \in [0, \pi]$  in (3), then  $\lambda_n = n^2$ ,  $n \geq 0$ .

**Theorem 7** (Hambardzumyan, 1929). *If*

$$\lambda_n = n^2, \quad n \geq 0,$$

*then*  $q(x) \equiv 0$ ,  $x \in [0, \pi]$ .

In general one spectrum is not enough to construct Sturm–Liouville problem since two different Sturm–Liouville boundary problems may have the same spectrum. For example,  $\lambda_n = n^2$ ,  $n \geq 0$  is the spectrum of the following boundary value problems:

$$\begin{aligned} -y'' &= \lambda y, \quad y'(0) = y'(\pi) = 0, \\ -y'' + \frac{2}{(1+x)^2}y &= \lambda y, \quad y'(0) + y(0) = 0, \quad y'(\pi) + \frac{1}{\pi+1}y(\pi) = 0. \end{aligned}$$

Consider two different Sturm–Liouville problems:

$$\begin{cases} -y'' + q(x) = \lambda y, & x \in [0, \pi], \\ y'(0) - hy(0) = 0, \\ y'(\pi) + Hy(\pi) = 0, \end{cases} \quad (4)$$

$$\begin{cases} -y'' + \tilde{q}(x) = \lambda y, & x \in [0, \pi], \\ y'(0) - \tilde{h}y(0) = 0, \\ y'(\pi) + \tilde{H}y(\pi) = 0, \end{cases} \quad (5)$$

where  $q(x), \tilde{q}(x) \in [0, \pi]$  are real continuous functions and  $h, H, \tilde{h}$  and  $\tilde{H}$  are finite real numbers.

We denote by  $\varphi(x, \lambda)$  and  $\tilde{\varphi}(x, \lambda)$  the solutions of the equations in (4) and (5) satisfying the following initial conditons:

$$\begin{aligned} \varphi(0, \lambda) &= 1, & \varphi'(0, \lambda) &= h, \\ \tilde{\varphi}(0, \lambda) &= 1, & \tilde{\varphi}'(0, \lambda) &= \tilde{h}, \end{aligned}$$

respectively. Then we have the spectral data formed by  $\{\lambda_n, \alpha_n\}_{n=0}^\infty$  and by  $\{\tilde{\lambda}_n, \tilde{\alpha}_n\}_{n=0}^\infty$  for each problem. So we arrive at another well known uniqueness theorem.

**Theorem 8** (Marchenko, 1977). *If*

$$\lambda_n = \tilde{\lambda}_n, \quad \alpha_n = \tilde{\alpha}_n, \quad n \geq 0,$$

*then the following equalities hold:*

$$h = \tilde{h}, \quad H = \tilde{H}, \quad q(x) = \tilde{q}(x), \quad x \in [0, \pi].$$

We state next the theorem of Borg (see [2]). Let us be given four different boundary value problems:

$$\begin{cases} -y'' + q(x) = \lambda y, & x \in [0, \pi], \\ y'(0) - hy(0) = 0, \\ y'(\pi) + Hy(\pi) = 0. \end{cases}$$

$$\begin{cases} -y'' + q(x) = \lambda y, & x \in [0, \pi], \\ y'(0) - h_1y(0) = 0, \\ y'(\pi) + Hy(\pi) = 0. \end{cases}$$

$$\begin{cases} -y'' + \tilde{q}(x) = \lambda y, & x \in [0, \pi], \\ y'(0) - \tilde{h}y(0) = 0, \\ y'(\pi) + \tilde{H}y(\pi) = 0. \end{cases}$$

$$\begin{cases} -y'' + \tilde{q}(x) = \lambda y, & x \in [0, \pi], \\ y'(0) - \tilde{h}_1y(0) = 0, \\ y'(\pi) + \tilde{H}y(\pi) = 0. \end{cases}$$

where  $q(x), \tilde{q}(x) \in [0, \pi]$  are real continuous functions and  $h, h_1, \tilde{h}, \tilde{h}_1, H$  and  $\tilde{H}$  are finite real numbers. We denote by  $\{\lambda_n\}_{n=0}^\infty, \{\mu_n\}_{n=0}^\infty, \{\tilde{\lambda}_n\}_{n=0}^\infty$  and  $\{\tilde{\mu}_n\}_{n=0}^\infty$  the eigenvalues of these problems.

**Theorem 9** (G. Borg, [2]). *If*

$$\lambda_n = \tilde{\lambda}_n, \quad \mu_n = \tilde{\mu}_n, \quad n = 0, 1, 2, \dots,$$

*then*

$$q(x) = \tilde{q}(x), \quad h = \tilde{h}, \quad h_1 = \tilde{h}_1, \quad H = H_1.$$

## Gelfand–Levitan algorithm

In order to learn to solve an inverse problem we introduce some necessary information. Later on we use the following theorems.

**Theorem 10.** *For the spectral data  $\{\lambda_n\}_{n=0}^{\infty}$  and  $\{\alpha_n\}_{n=0}^{\infty}$  the following asymptotic representations hold:*

$$\begin{cases} \sqrt{\lambda_n} = n + \frac{\omega}{n\pi} + \frac{\chi_n}{n}, & \{\chi_n\} \in l_2, \\ \alpha_n = \frac{\pi}{2} + \frac{\chi_n}{n}, & \{\chi_n\} \in l_2, \\ \lambda_n \neq \lambda_m, \quad n \neq m, \quad \alpha_n > 0, & n = 0, 1, 2, \dots, \end{cases} \quad (6)$$

where

$$\alpha_n = \int_0^\pi \varphi^2(x, \lambda_n) dx, \quad n = 0, 1, 2, \dots$$

**Theorem 11.** *For the solution  $\varphi(x, \lambda)$  of the equation (1) satisfying the initial conditions  $\varphi(0, \lambda) = 1$ ,  $\varphi'(0, \lambda) = h$  the following integral representation holds:*

$$\varphi(x, \lambda) = \cos \sqrt{\lambda}x + \int_0^x K(x, t) \cos \sqrt{\lambda}t dt, \quad (7)$$

where  $K(x, t)$  is any real continuous function satisfying the condition:

$$K(x, x) = h + \frac{1}{2} \int_0^x q(t) dt.$$

We know that the spectral data of the problem (1)–(2) satisfies (6). Now we look for a way of solving an inverse problem that is given by two sequences satisfying the condition (6). By (7) we can see that if we get  $K(x, t)$  by the spectral data, then we can find  $q(x)$ ,  $h$  and  $H$  using the following formulas:

$$q(x) \equiv 2 \frac{d}{dx} K(x, x), \quad h = K(0, 0), \quad H = \omega - h - \frac{1}{2} \int_0^\pi q(t) dt. \quad (8)$$

This information is not enough to solve the problem, that is why we define the following function by spectral data so that we can get the function  $K(x, t)$ :

$$F(x, t) = \sum_{n=0}^{\infty} \left( \frac{\cos \sqrt{\lambda_n}x \cos \sqrt{\lambda_n}t}{\alpha_n} - \frac{\cos nx \cos nt}{\alpha_n^0} \right), \quad (9)$$

where

$$\alpha_n^0 = \begin{cases} \frac{\pi}{2}, & n > 0, \\ \pi, & n = 0. \end{cases}$$

The following theorem plays a crucial role in solving the inverse problem.

**Theorem 12.** *For each fixed  $x \in [0, \pi]$ ,  $K(x, t)$  satisfies the following linear integral equation:*

$$K(x, t) + F(x, t) + \int_0^x K(x, s)F(s, t) ds = 0, \quad 0 < t < x. \quad (10)$$

Equation (10) is called Gelfand–Levitan integral equation or the main integral equation of the inverse problem theory [4]. So, we have all we need to solve the inverse problem.

The algorithm for solving an inverse problem is as follows:

1. By the spectral data we define  $F(x, t)$  function using (9),
2. Inserting  $F(x, t)$  into (10) we get  $K(x, t)$ ,
3. Employing  $K(x, t)$  in (8) we finally obtain  $q(x)$ ,  $h$  and  $H$ .

**Example 13.** *Let's try to solve one inverse problem using Gelfand–Levitan algorithm.*

$$\text{Given } \lambda_n = n^2, \quad \alpha_n = \begin{cases} \frac{\pi}{2}, & n \geq 1, \\ \alpha_0^n, & n = 0, \end{cases} \quad \alpha_0^n = \begin{cases} \frac{\pi}{2}, & n \geq 0, \\ \pi, & n = 0, \end{cases} \quad \text{find } \{q(x), h, H\}.$$

According to the algorithm above we get:

1.  $F(x, t) = \frac{1}{\alpha_0} - \frac{1}{\pi} = a$  by (9).
2. Inserting  $F(x, t)$  into (10) we get  $K(x, t) = -\frac{a}{1+ax}$ .
3. Employing  $K(x, t)$  in (8) we finally obtain  $q(x) = \frac{2a^2}{(1+ax)^2}$ ,  $h = -a$ ,  $H = \frac{a}{1+a\pi} = \frac{a\alpha_0}{\pi}$ .

Then, the function  $\varphi(x, \lambda)$  has the following form:

$$\varphi(x, \lambda) = \cos \sqrt{\lambda}x - \frac{a}{1+ax} \frac{\text{sen } \sqrt{\lambda}x}{\sqrt{\lambda}}.$$

Finally, we construct the following boundary value problem:

$$\begin{cases} -y'' + \frac{2a^2}{(1+ax)^2}y = \lambda y, \\ y'(0) + ay(0) = 0, \\ y'(\pi) + \frac{a\alpha_0}{\pi}y(\pi) = 0. \end{cases}$$

The problem is completely solved.

## References

- [1] Hambardsumyan, V. A. (1929) *Über eine Frage der Eigenwerttheorie*, Zeitschrift für Physik.
- [2] Borg, G. (1946). *Eine Umkehrung der Sturm–Liouvilleschen Eigenwertaufgabe*, Acta Mathematica.
- [3] Marchenko, V. A. (1977). *Sturm–Liouville operators and their applications*, Naukova dumka.
- [4] Levitan, B. (1984). *Sturm–Liouville inverse problems*, Nauka.
- [5] Xasanov, A. (2011). *Introduction to the theory of Sturm–Liouville boundary value problems*, Fan.