

MEMORIAS DE LA XXVIII ESCUELA DE VERANO EN FÍSICA
JUNIO 21 - JULIO 2, 2021
FORMATO VIRTUAL

José Récamier

Rocío Jáuregui

Instituto de Ciencias Físicas

Instituto de Física

Universidad Nacional Autónoma de México

Esta edición fue preparada por el Instituto de Física y
el Instituto de Ciencias Físicas de la UNAM.

Primera edición electrónica: Enero de 2022

D.R. © 2022 UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO

Ciudad Universitaria, 04510, Ciudad de México
Instituto de Física
Instituto de Ciencias Físicas

Prohibida la reproducción parcial o total por cualquier medio sin autorización escrita de su
legítimo titular de derechos.

ISSN: 2594-2697

Hecho en México

Contenido

Contribuciones
Introducción
Agradecimientos
Profesores participantes
Alumnos participantes

• Contribuciones

Denis Boyer <i>Procesos de reinicio estocástico y sus aplicaciones</i>	14
Neil C Bruce <i>Medición precisa de la polarización de la luz</i>	29
Santiago F. Caballero Benítez <i>Materia cuántica en cavidades de alta reflectancia (Many-body CQED)</i>	47
<i>S. Castillo-López, R. Esquivel-Sirvent, G. Pirruccio y C. Villarreal</i> Fuerzas y calor a la nanoescala	54
<i>Víctor Contreras</i> Levitadores acústicos uniaxiales: explorando sus límites para mejorar su desempeño	60
<i>Sebastien Fromenteau</i> Galaxy-dark matter connection in Cosmology	67
<i>Guillermo Hinojosa Aguirre,</i> Los iones negativos como prueba crítica de la mecánica cuántica	76
<i>Rocío Jáuregui,</i> Sobre las propiedades mecánicas de la luz	84
<i>Merlyn Jaqueline Juárez Gutiérrez y W. Luis Mochán</i> Cálculo de propiedades ópticas de metamateriales	91

<i>Hugo Alberto Lara García</i> Semiconductores de baja dimensionalidad	127
<i>Roberto de J. León Montiel</i> Simulaciones cuánticas en sistemas electrónicos sintéticos	134
<i>Eugenio Ley Koo</i> Fundamentos de electromagnetismo y de mecánica cuántica (I-III)	139
<i>François Leyvraz</i> Percolación y sus variaciones: transiciones geométricas	146
<i>L. Medina Dozal, I Ramos Prieto y José Récamier</i> Evolución temporal aproximada de un sistema optomecánico forzado con un sistema de dos niveles dentro de la cavidad	158
<i>Rosario Paredes</i> Superfluidez en redes ópticas	171
<i>Pedro A Quinto Su</i> Micromáquinas, pinzas y burbujas	177
<i>Karla Récamier Méndez</i> La física en el análisis de moléculas y medicamentos biotecnológicos	186
<i>Penélope Rodríguez, Oliver Paz Borbón y Andrés Botello</i> Física a la nanoescala	193
<i>Genaro Toledo Sánchez</i> Del protón a los Quarks	222
<i>Mayo Villagrán Muñiz</i> Láseres: teoría y experimento	240

- **Introducción**

La XXVIII Escuela de Verano en Física fue organizada con apoyo del Posgrado en Ciencias Físicas por el Instituto de Física y el Instituto de Ciencias Físicas de la Universidad Nacional Autónoma de México (UNAM). La Escuela, realizada en el periodo comprendido entre el 21 de junio y el 2 de julio de 2021, se llevó a cabo de forma virtual por primera vez. Este esquema favoreció la participación de estudiantes de diversas entidades dentro de la República Mexicana y de entidades académicas hermanas localizadas en Colombia, Guatemala, Ecuador, Perú y Estados Unidos de América; asimismo impartieron conferencias y cursos investigadores ampliamente reconocidos de muy diversas dependencias cercanas a la Física dentro de la UNAM.

En esta edición de la Escuela se impartieron 9 cursos y 40 conferencias, cubriéndose un amplio espectro con temas como: sistemas cuánticos de muchos cuerpos, física de la nanoescala, átomos de Rydberg, partículas elementales, propiedades ópticas de metamateriales, biofísica molecular, física molecular y de plasmas, óptica cuántica, sistemas complejos, simulaciones cuánticas, biología cuántica, agujeros negros, láseres, complejidad, astronomía, interferometría, movimientos colectivos y biofísica, entre otros.

Las grabaciones de los cursos y conferencias impartidas en esta Escuela de Verano estarán a su disposición en https://bit.ly/Youtube_EVFUNAM

*Rocío Jáuregui, IF UNAM
José Récamier, ICF UNAM
Diciembre, 2021*

- **Agradecimientos**

Agradecemos los apoyos recibidos para la realización de esta escuela a la Universidad Nacional Autónoma de México a través de la Coordinación de la Investigación Científica, del Instituto de Física y del Instituto de Ciencias Físicas.

Agradecemos al Dr. César Leonardo Ordoñez, al Dr. Eduardo Espinoza Ávila, al Lic. Luis Novoa Sandoval, y a Daniel de Jesús Rosales Mendoza, por su colaboración en la organización de la infraestructura que nos permitió realizar esta escuela en formato virtual.

Asimismo, agradecemos a la Lic. Celia B. Herrera Zambrano, al Fis. Alberto García Ramírez y al Ing. Fernando Javier Martínez Mendoza quienes se encargaron del funcionamiento del sistema de videoconferencias.

Agradecemos a la Lic. Celia B. Herrera Zambrano por su excelente trabajo en la elaboración de este libro.

• **Profesores participantes**

- *Andrés Botello, Lauro Oliver Paz, Penélope Rodríguez, IF UNAM*
Física a la nanoescala
- *Rosario Paredes, Víctor Romero, Santiago Caballero, IF UNAM*
Algunos fenómenos novedosos en sistemas cuánticos de muchos cuerpos: superfluidez en redes ópticas, cruce BEC-BCS en gases atómicos de Fermi ultrafríos y materia cuántica en cavidades de alta-reflectancia
- *Asaf Paris Mandoki, IF UNAM*
Átomos de Rydberg
- *Genaro Toledo, IF UNAM*
Partículas elementales: del protón a los quarks
- *Raúl Esquivel, Giuseppe Pirruccio, Carlos Villarreal, IF UNAM*
Fuerzas y calor a la nanoescala
- *W. Luis Mochán, ICF UNAM*
Cálculo de propiedades ópticas de metamateriales
- *Iván Ortega, Humberto Saint Martin, ICF UNAM*
Biofísica molecular
- *Guillermo Hinojosa, Antonio Juárez, ICF UNAM*
Fundamentos y aplicaciones de Física molecular y de plasmas
- *Pablo Barberis, IIMAS UNAM*
Óptica cuántica
- *José Luis Ruvalcaba, IF UNAM*
Métodos no destructivos mediante espectroscopias aplicados a la arqueología y el arte
- *Denis Boyer, FC UNAM*
Procesos con reinicio estocástico y sus aplicaciones
- *Hugo Alberto Lara García, IF UNAM*
Nuevas tendencias en la sintonización de propiedades optoelectrónicas de semiconductores.
- *Gustavo Medina Tanco, ICN UNAM*
LINX: Desarrollo de tecnología espacial en física de astropartículas y derrama en aplicaciones sociales y comerciales

- *Pedro Antonio Quinto Su, ICN UNAM*
Micromáquinas, burbujas y pinzas ópticas
- *Roberto J. de León Montiel, ICN UNAM*
Simulaciones cuánticas en sistemas electrónicos sintéticos
- *Octavio Castaños, ICN UNAM*
Transiciones de fase en óptica cuántica
- *Erick Vázquez Jáuregui, IF UNAM*
Neutrinos, materia oscura y la naturaleza del Universo
- *Fernando Rojas, CNN UNAM*
Biología cuántica: Magnetorecepción y Fotosíntesis
- *Thomas Stegmann, ICF UNAM*
Óptica electrónica en nuevos materiales bi-dimensionales
- *Alejandro Reyes Coronado, FC UNAM*
Respuesta óptica de metasuperficies desordenadas y aplicaciones
- *Juan Carlos Hidalgo, ICF UNAM*
Agujeros negros primordiales: Bichos raros del universo
- *Mayo Villagrán, ICAT UNAM*
Aplicaciones de láseres pulsados de nano segundos
- *Neil C Bruce Davidson, ICAT UNAM*
Medición precisa de la polarización de la luz
- *François Leyvraz, ICF UNAM*
Percolación y sus variaciones: transiciones geométricas
- *Alejandro Frank, ICN UNAM*
Matemáticas, Física y Biología: la complejidad en busca de señales de la autoorganización y la vida
- *Eugenio Ley Koo, IF UNAM*
Fundamentos comunes de Electromagnetismo y Mecánica Cuántica
- *José Récamier, ICF UNAM*
Sistemas optomecánicos híbridos
- *Gustavo Martínez Mekler, ICF UNAM*
Rasgos de Complejidad en la Fecundación
- *Mathieu Hautefeuille, FC UNAM*
Desarrollo de modelos microfisiológicos relevantes para la biomedicina usando principios físicos y microfabricación

- *Victor Contreras, ICF UNAM*
Levitadores acústicos uniaxiales: explorando sus límites para superar sus limitaciones
- *William Lee Alardin, IA UNAM*
¿Qué pasa cuando chocan dos estrellas de neutrones?
- *Gloria Koenigsberger, ICF UNAM*
Los telescopios espaciales
- *Guillermo Hinojosa, ICF UNAM*
Iones negativos
- *Hernán Larralde, ICF UNAM*
Caminatas aleatorias
- *Luis Benet, ICF UNAM*
Aceleración no cero de Yarkovsky para Apophis
- *Remigio Cabrera, ICF UNAM*
Dinámica electrón-núcleo: procesos dependientes del tiempo
- *Gerardo García Naumis, IF UNAM*
Superconductividad y ecuaciones relativistas en materiales bidimensionales
- *Alberto Güijosa, PCF UNAM*
Posgrado en Ciencias Físicas
- *Alberto Güijosa, ICN UNAM*
Holografía, Entrelazamiento Cuántico y Gravedad
- *Gabriel Vázquez Torres, ICF UNAM*
Introducción a la espectroscopia atómica y molecular: aplicaciones atmosféricas y astronómicas.
- *Fernando Ramírez, ICN UNAM*
Métodos experimentales básicos de interferometría con ondas de materia
- *Sebastien Fromenteau, ICF UNAM*
Conexión galaxia-materia oscura en cosmología
- *Frédéric Masset, ICF UNAM*
Ecuaciones de la Magnetohidrodinámica ideal
- *Juan Valentín Escobar, IF UNAM*
Procesos dinámicos en superficies

- *Libertad Barrón Palos, IF UNAM*
Neutrones y violación de inversión temporal
- *Karla Récamier, LAMMB*
La física en el análisis de moléculas y medicamentos biotecnológicos
- *Rocío Jáuregui, IF UNAM*
Efectos mecánicos de la luz

Simbología:

<i>IF UNAM</i>	<i>Instituto de Física, UNAM</i>
<i>ICAT UNAM</i>	<i>Instituto de Ciencias Aplicadas, UNAM</i>
<i>ICN UNAM</i>	<i>Instituto de Ciencias Nucleares, UNAM</i>
<i>IA UNAM</i>	<i>Instituto de Astronomía, UNAM</i>
<i>LAMMB</i>	<i>Laboratorio de Análisis de Moléculas y Medicamentos Biotecnológicos</i>
<i>FC UNAM</i>	<i>Facultad de Ciencias, UNAM</i>
<i>PCF UNAM</i>	<i>Posgrado en Ciencias Físicas, UNAM</i>
<i>CNN UNAM</i>	<i>Centro de Nanociencias y Nanotecnología, UNAM</i>
<i>IIMAS UNAM</i>	<i>Instituto de Investigaciones en Matemáticas Aplicadas y en Sistemas, UNAM</i>

Alumnos participantes

1. *Aarón Rodríguez González Pacheco (FC UNAM)*
2. *Adán Castillo Guerrero (FC UNAM)*
3. *Adrián Bruce Nolasco Cabello (FC UNAM)*
4. *Aldair Hernández Cruz (FES-Cuatitlán UNAM)*
5. *Alejandra Dávila Rivera (Universidad Autónoma de Zacatecas)*
6. *Alejandro Alvarado Corzo (FC UNAM)*
7. *Alexis Paul Vera Tapia (Escuela Superior Politécnica de Chimborazo - Ecuador)*
8. *Ana Gabriela Fernández Morantes (Universidad de Cauca - Colombia)*
9. *Antonio Moreno Herrera (Universidad Autónoma de Chihuahua)*
10. *Aurora Mata Sánchez (FC UNAM)*
11. *Bianca Michelle Cisneros Figueroa (Facultad de Ciencias Físico Matemáticas, Universidad Autónoma de Coahuila)*
12. *Brian Zamora Martínez (Instituto Tecnológico de Celaya)*
13. *Carlos Antonio Beauregard León (University of the Ozarks- EUA)*
14. *Carmen Lizet Seminario Panta (Universidad Nacional de Piura-Perú)*
15. *César Enrique Terán Cisneros (Facultad de Ciencias Físico Matemáticas, Universidad Autónoma de Coahuila)*
16. *Clover Emmanuel de la Cruz (Facultad de Ciencias, Universidad Autónoma del Estado de Morelos)*
17. *Diana Carolina Méndez Luna (Facultad de Ciencias Físico Matemáticas, Universidad Autónoma de Coahuila)*
18. *Diana Itzel Flores Baeza (FC UNAM)*
19. *Diana Laura Quijada Ocampo (FC UNAM)*
20. *Diego Emilio Domínguez Tableros (Escuela Superior de Física y Matemáticas, Instituto Politécnico Nacional)*
21. *Edgar Muñoz Sánchez (FC UNAM)*

22. *Eduardo Oropeza (Universidad Autónoma de Chapingo)*
23. *Elizabeth Rubio Juárez (FC UNAM)*
24. *Emmanuel Luna Avilés (FC UNAM)*
25. *Erick Manuel Pineda Ríos (FC UNAM)*
26. *Erick Sebastián Navarrete Cruz (FC UNAM)*
27. *Erika Cecilia Carrillo Trejo (Universidad Autónoma de Zacatecas)*
28. *Francisco Barreto Basave (Facultad de Ciencias, Universidad Autónoma del Estado de Morelos)*
29. *Francisco Guadalupe Jiménez López (Facultad de Ciencias, Universidad Autónoma de Nuevo León)*
30. *Gabriel Rodríguez Guijarro (Universidad Autónoma de Zacatecas)*
31. *Gabriela Isabel Vera Garfias (Universidad Autónoma Metropolitana-Atcapozalco)*
32. *Gil Estéfano Rodríguez Rivera (Universidad Autónoma de Guanajuato)*
33. *Guillermo Segura Gómez (Universidad Autónoma de Guanajuato)*
34. *Guillermo Sierra Vargas (Universidad Autónoma de San Luis Potosí)*
35. *Isabel Lucía Constantino Preciado (Universidad Autónoma de Nuevo León)*
36. *Jair Othoniel Domínguez Godínez (Universidad Autónoma de Nuevo León)*
37. *Jesús Emmanuel Gómez Vargas (FC Universidad Autónoma del Estado de México)*
38. *Jorge Alejandro Ávalos Haidacher (Escuela Superior de Física y Matemáticas, Universidad de San Carlos - Guatemala)*
39. *José Andrés Evangelio Vergaray (Escuela Superior de Física y Matemáticas, Universidad de San Carlos - Guatemala)*
40. *José Carlos Joaquín Altamirano (FC UNAM)*
41. *José Carlos Orozco Álvarez (Universidad de Guadalajara)*
42. *José María Mondragón Álvarez (Facultad de Ciencias, Universidad Autónoma del Estado de Mexico)*

43. *Juan Augusto Cabrera Padilla (FC UNAM)*
44. *Litzy Lilián García Faustino (Universidad de las Américas, Puebla)*
45. *Lucero Juárez Pinedo (Universidad Autónoma de Zacatecas)*
46. *Manuel Fernando López Rodríguez (FC UNAM)*
47. *María Belen Amaguaña Guevara (Escuela Superior Politécnica de Chimborazo - Ecuador)*
48. *María Guadalupe Mendoza Granados (Universidad Autónoma de Guanajuato)*
49. *Mario César Uyoa López (FC UNAM)*
50. *Mauricio Vladimir Tipa Valdevieso (Universidad Autónoma Metropolitana)*
51. *Miguel Ángel Agama Flores (Facultad de Ciencias Físicas Universidad Nacional Mayor de San Marcos - Perú)*
52. *Nahúm Efrén Vázquez Espinosa (FC UNAM)*
53. *Natalia Baez de la Cruz (FC UNAM)*
54. *Natalia Edith Mejía Bautista (FC UNAM)*
55. *Pedro Pablo Rosario Vargas (Universidad Autónoma de San Luis Potosí)*
56. *Raúl Eduardo Santoy Flores (Universidad Autónoma de Nuevo León)*
57. *Salem Hernández Ríos (Universidad Autónoma de Chihuahua)*
58. *Sergio Eduardo Rodríguez Tinoco (FC UNAM)*
59. *Ximena de la Rosa Gómez (FC UNAM)*
60. *Zenaida Berenice Sánchez Méndez (FC UNAM)*

Procesos de reinicio estocástico y sus aplicaciones

Denis Boyer

Instituto de Física, UNAM

Noviembre 2021

1 Introducción

En la naturaleza abundan los procesos de búsqueda. Hablamos de búsqueda cuando una entidad (buscador) se mueve en un espacio con el propósito de encontrar otra entidad (blanco). La búsqueda es esencial en la vida: en los procesos de polinización, las partículas de polen vuelan en el aire hasta llegar (quizás) al pistilo de otra flor para su reproducción; los animales recorren su medio ambiente para encontrar fuentes de alimento. Los virus o los antígenos se difunden en el medio intercelular “en busca” de cierto tipo de células para amarrarse en unos receptores específicos de membrana, lo cual iniciará procesos importantes en dichas células. En contextos tecnológicos, los buscadores del Internet usan procesos que se difunden en redes complejas de páginas web, hasta encontrar las más relevantes. En la vida cotidiana también nos topamos con problemas de búsqueda, como encontrar las llaves o reconocer la cara de una persona conocida en una multitud. En la figura 1 mostramos un ejemplo de búsqueda visual.

Una pregunta central en el estudio de este tipo de problemas sería: ¿qué estrategia tendría que adoptar un buscador para encontrar rápidamente su objetivo? Como el buscador desconoce la ubicación de este último, podría escoger un movimiento determinista y sistemático (por ejemplo en forma de espiral alrededor de su punto de partida, como en la Fig. 1) o seguir una estrategia no tan ordenada, dando pasos al azar en todas las

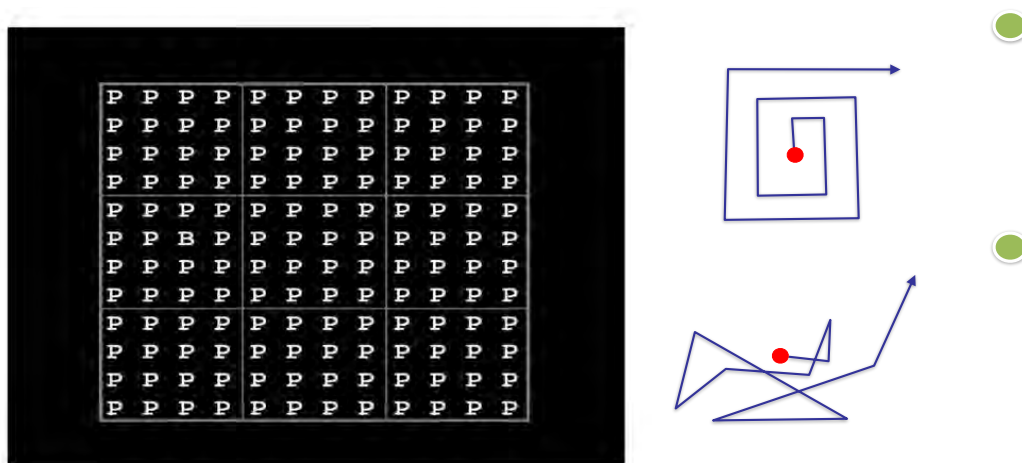


Figure 1: Búsqueda visual de una letra diferente de “P” y posibles estrategias de búsqueda (determinista, aleatoria).

direcciones. Por ser muy común, nos interesaremos en este texto a esta última clase de procesos, las búsquedas aleatorias, donde el azar y el ruido juegan un papel esencial. Después de recordar algunas propiedades básicas de las caminatas aleatorias, presentaremos los procesos de reinicio, que han atraído mucho interés en los últimos diez años en la comunidad de la física estadística. Veremos cómo el simple hecho de reiniciar un proceso de vez en cuando a su punto de partida puede mejorar considerablemente la eficiencia de su búsqueda. Presentaremos dos aplicaciones recientes de esta idea: las búsquedas en redes complejas y el movimiento de partículas brownianas en potenciales intermitentes.

2 Movimiento aleatorio

2.1 Difusión simple

Al principio del siglo XX, los matemáticos Karl Pearson y George Pólya fueron pioneros en el estudio de las caminatas aleatorias, un tipo de movimiento no determinista que iba a tener un número considerable de aplicaciones en muchas áreas de la ciencia. Consideremos una partícula que se mueve sobre una red unidimensional de sitios (ver figura 2), de tal manera que a cada paso de tiempo la partícula brinca a un sitio vecino de su posición actual. Puede escoger con igual probabilidad ($1/2$) ir al sitio de la izquierda o al de la derecha. Cada paso es independiente del anterior. Después de t pasos, se genera una trayectoria errática (es decir un proceso estocástico) como la que se muestra en la figura 2 hasta un tiempo $t \gg 1$. La posición $X(t)$ de la partícula después de exactamente t pasos es una variable aleatoria: fluctúa de una trayectoria a la otra. Podemos introducir la probabilidad $P_t(n)$ que el caminante se encuentre en el sitio n después de t pasos, o $P_t(n) = \text{Proba}[X(t) = n]$, con $n \in \mathbb{Z}$. Esta función de dos variables satisface la relación de recurrencia:

$$P_t(n) = \frac{1}{2}P_{t-1}(n-1) + \frac{1}{2}P_{t-1}(n+1), \quad (1)$$

la cual traduce el hecho que para estar en el sitio n al tiempo t , el caminante estaba justo al tiempo anterior $t-1$ en el sitio vecino $n-1$ y dió un paso a la derecha (lo cual ocurre con probabilidad $1/2$), o estaba en el sitio $n+1$ y dió un paso a la izquierda (de probabilidad también $1/2$). Tomando el límite continuo ($n \gg 1, t \gg 1$), se expande (1) en series de Taylor para llegar a:

$$\frac{\partial P}{\partial t} = D \frac{\partial^2 P}{\partial n^2}, \quad (2)$$

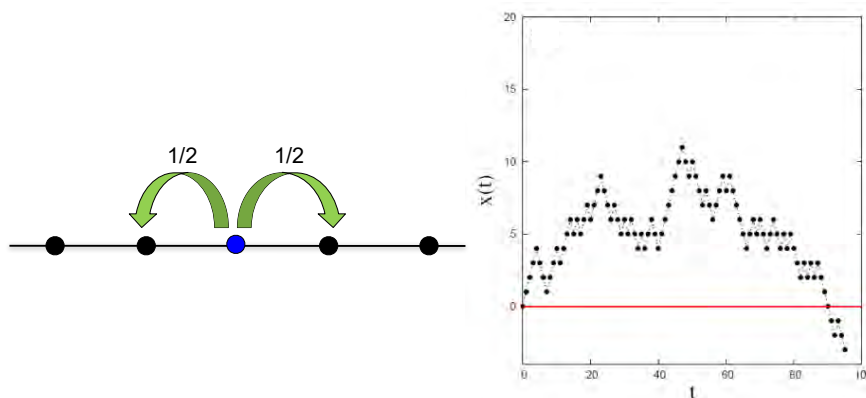


Figure 2: Caminante aleatorio en tiempo discreto en una red unidimensional y un ejemplo de trayectoria generada por t pasos.

con $D = 1/2$ aquí. La ecuación (2) es la ecuación de difusión, con constante de difusión D , una de las ecuaciones con derivadas parciales más conocida de la física. Es un caso particular de ecuaciones de Fokker-Planck, que son ecuaciones que describen la evolución en el tiempo de una distribución de probabilidad de una variable aleatoria. Esa variable puede tomar valores arbitrarios, las posiciones n aquí. En el caso de tener muchas partículas que difunden sin interactuar entre sí, se puede usar la misma ecuación (2) para describir la evolución de su concentración.

Si la partícula parte de la posición $n = 0$ a $t = 0$, la solución de (2) está dada para $t > 0$ por una función gaussiana:

$$P_t(n) = \frac{1}{\sqrt{4\pi Dt}} \exp\left(-\frac{n^2}{4Dt}\right). \quad (3)$$

Vemos de esta función es par en n : por la simetría del problema, es equiprobable alcanzar la posición n o la posición $-n$ al tiempo t , por lo que el desplazamiento neto de la partícula es cero, es decir $\langle X(t) \rangle = \sum_{n=-\infty}^{\infty} n P_t(n) = 0$ (nótese que en el continuo las sumas se sustituyen por integrales). No existe una dirección preferencial que genere un flujo hacia un lado. Sin embargo, la distribución Gaussiana tiene un ancho, el denominador $4Dt$ en la exponencial de (3), que crece con t . Se puede mostrar que las fluctuaciones de la posición alrededor de su media están dadas por

$$\langle X^2(t) \rangle = \sum_{n=-\infty}^{\infty} n^2 P_t(n) = 2Dt. \quad (4)$$

Se trata de la bien conocida ley de Smolushovsky-Einstein para la difusión, que se aplicó por primera vez al movimiento browniano de partículas microscópicas en suspensión en un fluido. $\langle X^2(t) \rangle$ representa el desplazamiento cuadrático medio. Típicamente, después de t pasos la partícula se encuentra a una distancia $\sqrt{\langle X^2(t) \rangle} \propto \sqrt{t}$ de su punto de partida.

2.2 Tiempos de primer arribo

Se pueden abordar los procesos difusivos desde un punto de vista alterno, menos conocido tal vez que las consideraciones anteriores, pero muy relevante en el contexto de los procesos de búsqueda. Nos podemos hacer una pregunta diferente: ¿cuánto tiempo va a tardar la partícula difusiva en alcanzar por primera vez una posición dada? Supongamos que la partícula parte de una posición x_0 , como se indica en la figura 3-arriba, y que un blanco se encuentra en el origen. Después de un cierto tiempo t_1 , la partícula alcanza el blanco *por primera vez*. Por lo tanto, a $t = t_1$, la búsqueda concluye (si suponemos que la reacción entre ambas partes se da al primer encuentro).

El tiempo t_1 se denomina el tiempo de primer arribo y es una variable aleatoria: fluctúa de una trayectoria a la otra, puesto que estas son estocásticas. Hay varias preguntas que nos podemos hacer: ¿cuál será el tiempo promedio de primer arribo o de búsqueda, $\langle t_1 \rangle$? Aquí, el promedio $\langle \cdot \rangle$ de una cantidad se hace sobre todas las posibles trayectorias que podría recorrer la partícula que parte de x_0 . El tiempo promedio de primer arribo (MFPT, la abreviación en inglés de ‘mean first passage time’) obviamente depende de la distancia inicial x_0 al blanco y del tipo de movimiento que sigue la partícula (no tiene que describir necesariamente una caminata aleatoria simple). Luego, nos podemos también preguntar: a x_0 fijo, ¿se puede minimizar $\langle t_1 \rangle$, es decir, encontrar estrategias de búsqueda más eficientes que otras?

Otra cantidad de interés y que incluso da más información que $\langle t_1 \rangle$, es la función de distribución completa de la variable t_1 , notada como $F(x_0, t)$. De manera coloquial,

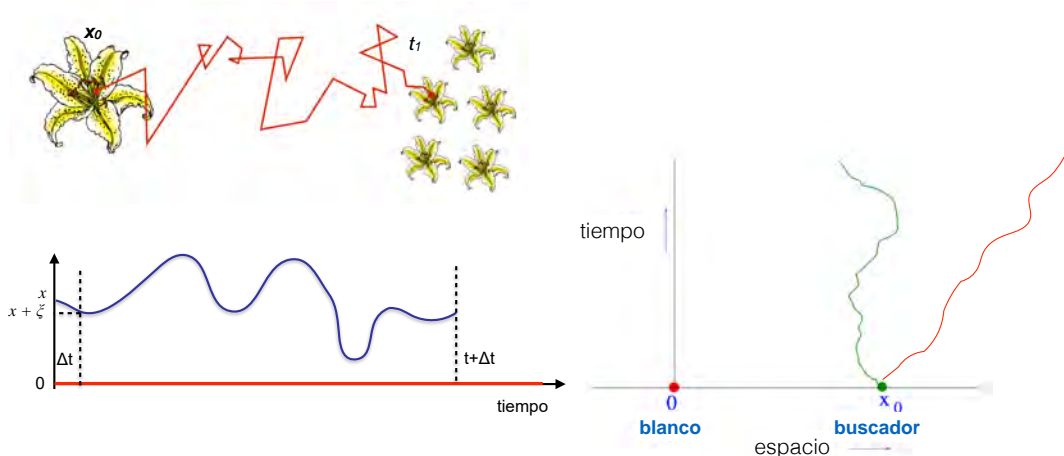


Figure 3: Arriba: un buscador (*e.g.*, una partícula de polen) sale de x_0 y llega por primera vez a un blanco al tiempo t_1 . Derecha: movimiento browniano unidimensional con un blanco ubicado en el origen. Se muestran dos trayectorias, donde una se aleja del objetivo. Izquierda-abajo: Un caminante sale de x a $t = 0$ y todavía no ha encontrado el blanco (en 0) al tiempo $t + \Delta t$.

$F(x_0, t)$ es la probabilidad que t_1 sea igual a t . Por ejemplo, si salimos en la noche en una ciudad desconocida, el tiempo que vamos a tomar para encontrar una tienda abierta puede ser de $t = 5$ minutos o $t = 1$ hora, o cualquier otro tiempo. Cada uno de los tiempos posibles de búsqueda tiene cierta probabilidad de ocurrir, dada por $F(x_0, t)$. También podemos introducir la probabilidad de ‘sobrevivencia’, $Q(x_0, t)$, que es la probabilidad de aún no haber encontrado el blanco al tiempo t . Se puede derivar fácilmente una relación general entre Q y F , válida para cualquier proceso:

$$F(x_0, t) = -\frac{\partial Q(x_0, t)}{\partial t}. \quad (5)$$

Ahora podemos considerar el problema de búsqueda aleatoria más simple, tal vez, ilustrado por la figura 3-derecha: consideremos un movimiento browniano, es decir, el límite continuo de la caminata aleatoria de la figura 2, donde tanto el espacio entre los sitios como el tiempo de cada salto tienden a cero. La partícula sale a $t = 0$ de x_0 y se difunde sobre una línea infinita en una dimensión espacial, con una constante de difusión D . Un blanco absorbente está ubicado en el origen. La probabilidad de supervivencia de la partícula se puede calcular exactamente,

$$Q(x_0, t) = \operatorname{erf}(|x_0|/\sqrt{4Dt}) \sim t^{-1/2} \text{ a } t \text{ grande}, \quad (6)$$

donde $\operatorname{erf}(z)$ es la función error, definida como $\operatorname{erf}(z) = \frac{2}{\sqrt{\pi}} \int_0^z dy e^{-y^2}$. Este resultado nos indica, primero, que Q decae a 0 cuando $t \rightarrow \infty$: todas las trayectorias terminan por encontrar el blanco. Sin embargo, el decaimiento es *muy lento*, dado que sigue una ley de potencia inversa a tiempos largos ($\propto 1/\sqrt{t}$). Ese tipo de función carece de escala temporal característica y por lo tanto se distingue drásticamente de un decaimiento exponencial, por ejemplo, como en el decaimiento radioactivo de los núcleos. Usando (5), deducimos

$$F(x_0, t) = \frac{|x_0|}{\sqrt{4\pi Dt^3}} \exp[-x_0^2/(4Dt)] \rightarrow t^{-3/2} \text{ a } t \text{ grande}, \quad (7)$$

que se conoce como la distribución de Lévy-Smirnov. Las trayectorias que más se tardan en encontrar el blanco (t grande) son menos frecuentes. En otras palabras, las excursiones

de larga duración t representan una fracción cada vez más pequeña, proporcional a $t^{-3/2}$, del total. Esta distribución de probabilidad es ‘libre de escala’ porque tiene una cola en ley de potencia. Tal decaimiento es lento y trae algunas consecuencias inesperadas. El tiempo promedio de búsqueda se obtiene de la definición del valor esperado $\sum_t t \text{Prob}[t]$ de una variable aleatoria, o $\langle t_1 \rangle = \int_0^\infty dt t F(x_0, t)$ en tiempo continuo aquí. En esta integral, se suma una función que decae como $t^{-1/2}$, por lo tanto

$$\langle t_1 \rangle = \infty. \quad (8)$$

¡La partícula se tarda un tiempo infinito para alcanzar el blanco! Este resultado puede parecer sorprendente y lo es. En realidad, la gran mayoría de las trayectorias encuentra el blanco en un tiempo corto, pero algunas pocas se tardan tanto que hacen divergir el promedio. Se trata de trayectorias, como la mostrada en la figura 3-derecha, que dan grandes excursiones del lado opuesto a donde se ubica el blanco.

Para demostrar el resultado (6) existen muchos métodos. Uno de los más sencillos y flexibles para generalizaciones a otros casos, es el de la ecuación ‘backward’ de Fokker-Planck. Consideremos en la figura 3-izquierda (abajo) una partícula que sale de una posición $x > 0$ a $t = 0$, en una dimensión espacial (la generalización a dimensiones superiores es sencilla). La probabilidad que la partícula no haya encontrado el blanco hasta el tiempo $t + \Delta t$ (es decir, haya permanecido por encima del eje horizontal $y = 0$) es $Q(x, t + \Delta t)$ por definición. Escogimos $\Delta t \ll t$ y descomponemos el intervalo de tiempo en dos: $[0, \Delta t]$ al inicio, seguido de otro de longitud t . Al tiempo Δt , la partícula se desplazó un poco de su posición inicial debido a la difusión y se encuentra en una nueva posición $x + \xi$, donde $|\xi| \ll x$. Como la difusión es un proceso sin memoria (o markoviano, o con pasos independientes) e invariante por translación temporal podemos considerar $x + \xi$ como un nuevo punto de partida. Entonces $Q(x, t + \Delta t)$ es igual a la probabilidad de no encontrar el blanco después de un tiempo t , partiendo de $x + \xi$:

$$Q(x, t + \Delta t) = \langle Q(x + \xi, t) \rangle_\xi. \quad (9)$$

En (9), hemos promediado sobre todos los desplazamientos ξ que puede dar la partícula durante el primer intervalo Δt . Expandimos esta ecuación de cada lado en series de Taylor,

$$Q(x, t) + \Delta t \frac{\partial Q(x, t)}{\partial t} + \dots = Q(x, t) + 0 + \frac{1}{2} \langle \xi^2 \rangle \frac{\partial^2 Q(x, t)}{\partial x^2} + \dots, \quad (10)$$

donde se ha usado el hecho que el término de primer orden $\langle \xi \rangle$ se anula por simetría: como en el caso de la caminata aleatoria, el movimiento browniano no tiene dirección preferencial. Por otra parte, usando la ley de Smolushovsky-Einstein (4) deducimos que $\langle \xi^2 \rangle = 2D\Delta t$. Después de agrupar los términos de orden Δt , llegamos a

$$\frac{\partial Q}{\partial t} = D \frac{\partial^2 Q}{\partial x^2}. \quad (11)$$

Entonces, Q satisface una ecuación de difusión, pero de un tipo peculiar: la variable x no es la posición ocupada al instante t , como acostumbramos, sino la *posición inicial*. Por eso se llama una ecuación ‘backward’ de Fokker-Planck [por oposición a una ecuación ‘forward’, como (2)]. Además $0 \leq Q \leq 1$ es una probabilidad y no una concentración o una densidad de probabilidad como en (2). Esta probabilidad satisface una condición inicial y una condición de frontera:

$$Q(x, t = 0) = 1, \quad x \neq 0 \quad (12)$$

$$Q(x = 0, t) = 0, \quad t > 0, \quad (13)$$

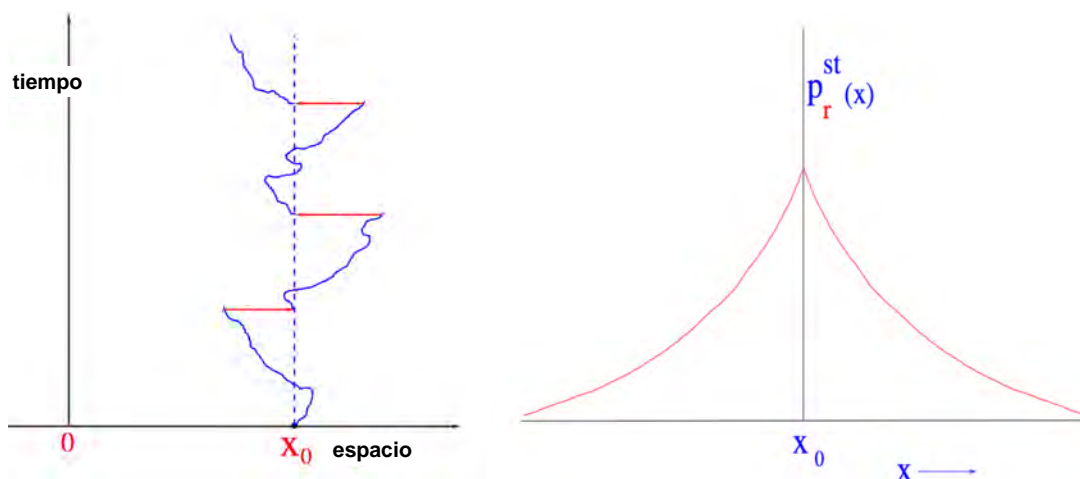


Figure 4: Izq.: Difusión con reinicio estocástico. Der.: distribución de probabilidad a $t \rightarrow \infty$.

que traducen el hecho que inicialmente la partícula tiene probabilidad uno de no haber encontrado el blanco, excepto si parte del blanco, justamente, donde la búsqueda termina inmediatamente. Si se resuelve (11) con (12)-(13), se obtiene (6) y las consecuencias que hemos discutido.

3 Difusión con reinicio estocástico

El resultado infinito (8) nos enseña que la difusión simple no representa una estrategia tan buena para encontrar un objetivo en un dominio no-acotado. En el 2011, Martin Evans y Satya Majumdar imaginaron una clase de procesos de búsqueda que sí se conclúan en un tiempo promedio finito: los procesos con reinicio estocástico ('stochastic resetting'). Consideremos por ejemplo una difusión simple con un ingrediente adicional: durante cualquier intervalo de tiempo dt , la partícula difusiva se puede reiniciar (o reubicar) a una posición fija con probabilidad $r dt$. Un nuevo parámetro interviene en el problema, la tasa r de reinicio¹. La figura 4-izq. muestra un ejemplo en $1d$, donde el sitio de reubicación es la posición inicial x_0 . Por simplicidad, se supone que las reubicaciones son instantáneas (se observan varias en la figura, marcadas por líneas horizontales).

Los procesos de reinicio tienen numerosas aplicaciones. Pueden describir el movimiento de un animal que regresa de vez en cuando a su nido o madriguera. O la movilidad humana. En ciencias de la computación, se usan algoritmos de reinicio para encontrar soluciones numéricas a problemas difíciles como el plegamiento de las proteínas o de optimización combinatoria: si volvemos a empezar de cero una búsqueda, podemos sacar un sistema de un estado poco óptimo y darle oportunidad de explorar otras regiones de su espacio-fase. También ciertas reacciones enzimáticas se pueden describir como procesos de reinicio (ver más adelante la figura 6-izq.). Otro ejemplo es la evolución de los genomas, que se da a través de pequeñas mutaciones azarosas en los pares de bases del ADN y que producen una lenta difusión en el espacio genómico. Sin embargo, a veces ayuda tener cambios un poco más 'violentos', tal como la duplicación de un golpe de todo un genoma, lo cual puede facilitar la búsqueda de estados más adaptados o de mayor 'fitness'. Un ejemplo más cotidiano es el duplicado de las llaves. Después de hacer copias de copias de copias, etc..., una llave acumula errores (difusión) a tal punto que llega un momento donde ya no abre: hay que recurrir al original, es decir reiniciar desde el origen otra vez.

¹Es una probabilidad por unidad de tiempo.

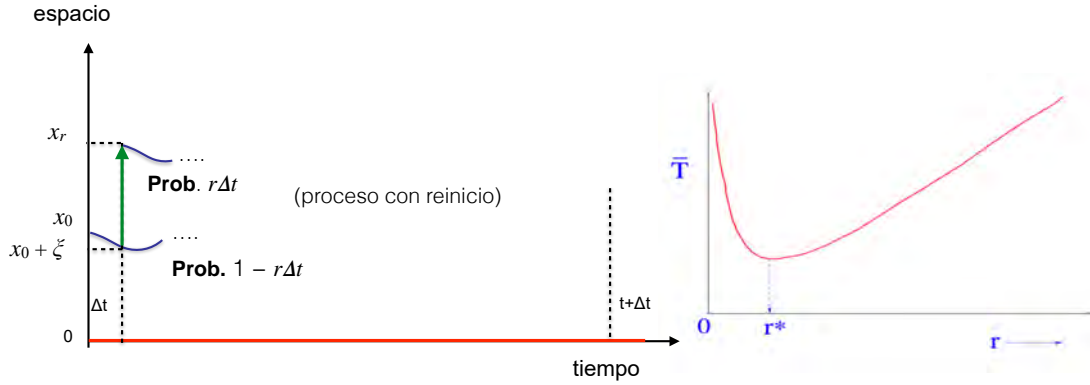


Figure 5: Izquierda: La partícula saliendo de x_0 se reubica con probabilidad $r\Delta t$ a x_r durante $[0, \Delta t]$, antes de seguir su camino. Sino, alcanza la posición $x_0 + \xi$. Derecha: tiempo promedio de primer arribo al origen como función de r .

Nuevamente, podemos analizar un problema simple, el movimiento browniano en $1d$ sujeto a reinicio, el primer caso considerado por Evans y Majumdar. El reinicio provoca dos cambios importantes, entre otros: la existencia de una distribución estacionaria y tiempos de primer arribo finitos. La primera propiedad se refiere a la distribución de la partícula sobre la línea infinita, sin blanco absorbente. Llamemos $p(x, t|x_0)$ la densidad de probabilidad que la partícula partiendo de x_0 a $t = 0$ se encuentre en x al tiempo t . La ecuación de la difusión (2) se modifica por la adición de dos términos:

$$\frac{\partial p(x, t|x_0)}{\partial t} = D \frac{\partial^2 p(x, t|x_0)}{\partial x^2} - rp(x, t|x_0) + r\delta(x - x_0), \quad (14)$$

que indican que la partícula desaparece de su posición actual a tasa r , mientras reaparece en x_0 con la misma tasa. Contrariamente a la solución gaussiana (3) de la ecuación de la difusión, que tiende a 0 para todo x cuando $t \rightarrow \infty$, la ecuación (14) admite una solución estacionaria no-nula, $p^{\text{st}}(x)$, que se alcanza a tiempos largos. Igualando el término de la izquierda de (14) a 0, se obtiene una distribución exponencial independiente del tiempo

$$p^{\text{st}}(x) = \frac{\alpha_0}{2} \exp(-\alpha_0|x - x_0|), \quad \text{con } \alpha_0 = \sqrt{r/D}, \quad (15)$$

que se grafica en la figura 4-derecha. Intuitivamente, la partícula nunca puede alejarse mucho de su punto de partida, puesto que el reinicio la trae constantemente ahí. Desde un punto de vista termodinámico, este proceso está *fuera de equilibrio* porque rompe el balance detallado: en un intervalo de tiempo dt , la partícula puede ir de cualquier x a x_0 , pero no puede ir de x_0 a cualquier x . Esto implica que haya flujos de probabilidad, aún en el estado estacionario. Por lo tanto, se habla de *estados estacionarios de no-equilibrio*.

Para estudiar las propiedades de primer arribo al origen, modificamos nuestro esquema de la figura 3 al de la figura 5-izq. La partícula parte de x_0 y puede ahora reubicarse en una posición x_r fija (la cual, en el caso general, puede ser igual a x_0 o no). Durante el intervalo $[0, \Delta t]$, sea la partícula no se reubica y se difunde hasta $x_0 + \xi$ (con probabilidad $1 - r\Delta t$), o se reubica en x_r (con la probabilidad $r\Delta t$). Entonces, la relación (9) para la probabilidad de sobrevivencia se vuelve

$$Q(x_0, t + \Delta t) = (1 - r\Delta t)\langle Q(x_0 + \xi, t) \rangle_\xi + r\Delta t Q(x_r, t). \quad (16)$$

Después de expandir en series de Taylor, se obtiene a orden Δt una ecuación ‘backward’

de Fokker-Planck:

$$\frac{\partial Q(x_0, t)}{\partial t} = D \frac{\partial^2 Q(x_0, t)}{\partial x_0^2} - rQ(x_0, t) + rQ(x_r, t). \quad (17)$$

[Aquí, $Q(x_0, t)$ depende implícitamente de la posición de reinicio x_r .] Existen técnicas para resolver (17) hasta cierto punto, pero nos podemos concentrar sobre el tiempo promedio de primer arribo $\langle t_1 \rangle(x_0)$, dado por $\langle t_1 \rangle(x_0) = \int_0^\infty dt t F(x_0, t) = \int_0^\infty dt Q(x_0, t)$. El último paso lo obtenemos de usar (5) y de integrar por partes. Integrando (17) en t de 0 a ∞ y usando la condición $Q(x_0, t = 0) = 1$ obtenemos una ecuación diferencial para $\langle t_1 \rangle(x_0)$,

$$D \frac{\partial^2 \langle t_1 \rangle(x_0)}{\partial x_0^2} - r \langle t_1 \rangle(x_0) = -1 - r \langle t_1 \rangle(x_r). \quad (18)$$

Considerando $\langle t_1 \rangle(x_r)$ como una constante, se resuelve en x_0 con la condición absorbente $\langle t_1 \rangle(0) = 0$ y de no-divergencia en $x_0 \rightarrow \infty$. Luego, $\langle t_1 \rangle(x_r)$ se determina por auto-consistencia y se obtiene, para el reinicio al punto de partida ($x_r = x_0$):

$$\langle t_1 \rangle(x_0) = \frac{1}{r} \left[\exp(x_0 \sqrt{r/D}) - 1 \right]. \quad (19)$$

Este resultado se representa en la figura 5-derecha. Ahora $\langle t_1 \rangle$ es finito para todo $r > 0$: el reinicio elimina esas trayectorias largas que causaban la divergencia del tiempo de búsqueda. Además, $\langle t_1 \rangle$ tiene variaciones no-monótonas con r y es *mínimo* para un cierto valor r^* . Existe entonces una tasa óptima de reinicio, para la cual la búsqueda es la más breve, en promedio. Se obtiene $r^* = z^* D/x_0^2$ con $z^* = 2.5396\dots$. Cuando $r \rightarrow 0$, recuperamos el movimiento browniano simple con $\langle t_1 \rangle \rightarrow \infty$. Cuando $r \rightarrow \infty$, también $\langle t_1 \rangle \rightarrow \infty$: el caminante siempre se queda muy cerca de su punto de partida y no puede alcanzar el origen.

4 Reinicio estocástico de caminatas aleatorias en redes

Vimos que el reinicio puede optimizar una búsqueda aleatoria en sistemas de baja dimensión espacial. Ahora ¿qué pasaría en un espacio más complejo de búsqueda, como una red arbitraria? Las redes están en todas partes y las caminatas en redes (con varios grados de complejidad) describen muchos fenómenos. En la figura 6 mostramos unos ejemplos: una secuencia de estados intermedios de una reacción química enzimática, donde pueden ocurrir retrocesos debido a imperfecciones; la movilidad de una persona visitando varios lugares que le son familiares en una ciudad; la búsqueda de páginas web en el World Wide Web, llevada a cabo por el algoritmo PageRank empleado por Google. Otro fenómeno de difusión en redes complejas es la propagación de enfermedades a través de la circulación de personas por la red mundial de transporte aéreo, con los posibles riesgos de pandemia asociados que conocemos.

Consideremos el problema planteado por el diagrama de la figura 6c. Un caminante (digamos, una turista) camina por una red arbitraria (digamos, de calles) y a cada paso de tiempo puede con probabilidad $1 - \gamma$ dar un paso aleatorio a un sitio vecino o con la probabilidad complementaria γ reiniciar desde un sitio r (su hotel). Si elige lo primero, w_{lm} denota la probabilidad de ir al sitio m desde el sitio ocupado l .

Nos podemos preguntar cómo el reinicio va a afectar la exploración de la red por la turista. Es de interés calcular la probabilidad $P_{ij}(t)$ de ocupar el sitio j después de

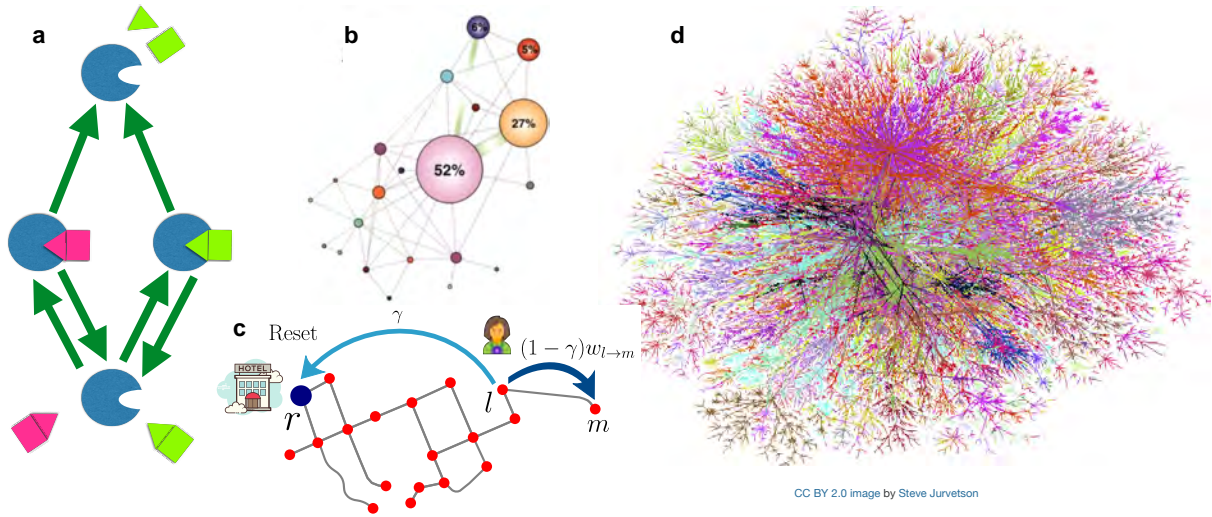


Figure 6: Caminatas en redes. **a)** Una red pequeña de reacción enzimática, con posibles regresos al estado inicial debido al rompimiento de un amarre físico [adaptado de Pal & Prasad, *Phys. Rev. Res.* **1**, 032001 (2019)]. **b)** Red de movilidad urbana de una persona a lo largo de 6 meses [tomado de C. Song *et al.*, *Science* **327**, 1018 (2010)]. **c)** Modelo de caminante en redes con regreso al sitio r (el ‘hotel’) con probabilidad γ y pasos aleatorios con probabilidad $1 - \gamma$ [Tomado de Riascos *et al.* 2020, ver Referencias]. **d)** Un mapa del Internet.

exactamente t pasos, partiendo de i . O el tiempo promedio de primer arribo a j desde i , $\langle T_{ij} \rangle$. Otra una cantidad que mide la capacidad de un caminante a explorar toda una red finita de N nodos es el *tiempo global* de primer arribo partiendo desde el sitio i , representado por un promedio sobre j :

$$T_i = \frac{1}{N} \sum_{j=1}^N \langle T_{ij} \rangle. \quad (20)$$

Además del reinicio, nos podemos preguntar cómo estas cantidades van a depender de la arquitectura de la red en cuestión. Por ejemplo, muchas redes tienen comunidades, es decir, grupos de nodos más conectados entre sí que con los otros nodos. ¿Cómo escoger el sitio de reinicio? También, ¿cuál va a ser el efecto del desorden, de las irregularidades en las conexiones?

Para definir una red arbitraria de N nodos, introducimos la matriz de adyacencia, de tamaño $N \times N$ y con elementos $A_{ij} = A_{ji} = 1$ si los nodos i y j están conectados y $A_{ij} = A_{ji} = 0$ sino. El grado k_i del nodo i se define como el número de conexiones que tiene, $k_i = \sum_{l=1}^N A_{il}$. Supongamos que la red no está desconectada en varias partes y que una caminata aleatoria recorre los nodos pasando por los enlaces existentes. La probabilidad de pasar del nodo l al nodo m en una unidad de tiempo está dada por $w_{lm} = A_{lm}/k_l$. Todos los vecinos de l tienen la misma probabilidad $1/k_l$ de ser visitados desde l , lo cual generaliza el caso de la figura 2. Notamos \mathbf{W} la matriz con elementos w_{lm} , que no es necesariamente simétrica. Según el modelo descrito arriba y en la figura 6c, la ecuación de evolución para $P_{ij}(t)$, o *ecuación maestra*, está dada por

$$P_{ij}(t+1) = (1 - \gamma) \sum_{l=1}^N P_{il}(t) w_{lj} + \gamma \delta_{rj}. \quad (21)$$

Esta ecuación generaliza (1) y agrega la reubicación directa al nodo r , que ocurre con

probabilidad γ . De manera más compacta, se puede re-escribir como

$$P_{ij}(t+1) = \sum_{l=1}^N P_{il}(t)\pi_{lj}, \quad (22)$$

donde $\pi_{lm} \equiv (1-\gamma)w_{lm} + \gamma\delta_{rm}$. La matriz $\mathbf{\Pi}$ con elementos π_{lm} es la matriz de transición del proceso. Vectorialmente, (22) se puede escribir como $\vec{P}(t+1)^T = \vec{P}(t)^T \mathbf{\Pi}$, donde $\vec{P}(t)^T$ es la transpuesta del vector de probabilidad de ocupación $\vec{P}(t)$, cuya j -ésima componente es $P_{ij}(t)$. Por iteración, la solución formal al tiempo t resulta de aplicar t veces la matriz $\mathbf{\Pi}$ a la condición inicial $P_{ij}(0) = \delta_{ij}$. En notaciones de Dirac, tenemos

$$P_{ij}(t) = \langle i | \mathbf{\Pi}^t | j \rangle, \quad (23)$$

donde $|i\rangle$ es el vector con entradas 0 excepto la i -ésima (que vale 1) y $\langle i|$ su transpuesta. Esta solución general se puede explicitar fácilmente si conocemos los eigenvalores y eigenvectores de $\mathbf{\Pi}$. Desde este enfoque, nuestro problema (y cualquier problema de difusión en una red) se reduce entonces a un problema de diagonalización de matrices. La matriz \mathbf{W} introducida anteriormente describe la caminata en la red *sin* reinicio, puesto que $\mathbf{\Pi}_{\gamma=0} = \mathbf{W}$. Supongamos que sepamos diagonalizar \mathbf{W} (analítica o numéricamente). Llamemos λ_l (con $l = 1, \dots, N$) los eigenvalores de \mathbf{W} , además de $|\phi_l\rangle$ y $\langle \bar{\phi}_l|$ sus eigenvectores derechos e izquierdos, respectivamente. Es decir,

$$\mathbf{W} |\phi_l\rangle = \lambda_l |\phi_l\rangle \quad \text{y} \quad \langle \bar{\phi}_l | \mathbf{W} = \lambda_l \langle \bar{\phi}_l|. \quad (24)$$

Afortunadamente, se pueden determinar los eigenvectores/valores de la matriz $\mathbf{\Pi}$ en función de $\{\lambda_l, |\phi_l\rangle, \langle \bar{\phi}_l|\}$, de γ y del sitio de reinicio $|r\rangle$.² Entonces, podemos obtener de (23) la probabilidad de ocupación a cualquier tiempo en función de estas mismas cantidades. Cuando $t \rightarrow \infty$, solamente el mayor eigenvalor, que es igual a 1, no decae a 0 y nos conduce a la probabilidad de ocupación estacionaria con reinicio al sitio r :

$$P_j^\infty(r) = \frac{k_j}{\sum_{m=1}^N k_m} + \gamma \sum_{l=2}^N \frac{\langle r | \phi_l \rangle \langle \bar{\phi}_l | j \rangle}{1 - (1-\gamma)\lambda_l}. \quad (25)$$

Para obtener el tiempo de primer arribo, podemos utilizar otro método que la ecuación ‘backward’ de Fokker-Planck. Se llama el método de la convolución y también funciona con cualquier proceso sin memoria. Consiste en darse cuenta que si el sitio j está ocupado al tiempo t , lo cual ocurre con probabilidad $P_{ij}(t)$, es que el caminante llegó ahí por primera vez en algún tiempo $t' \leq t$, para luego ocupar j nuevamente después de un lapso $t - t'$. Matemáticamente, se traduce como,

$$P_{ij}(t) = \delta_{t0}\delta_{ij} + \sum_{t'=0}^t P_{jj}(t-t')F_{ij}(t'), \quad (26)$$

donde $\delta_{t0}\delta_{ij}$ incorpora la condición inicial y $F_{ij}(t')$ es la probabilidad de alcanzar j por primera vez al tiempo t' desde i . [$F_{ij}(t)$ es similar a $F(x_0, t)$ en la ecuación (5).] En otras palabras, P_{ij} es una convolución entre F_{ij} y P_{jj} . Como conocemos $P_{ij}(t)$ y $P_{jj}(t)$, podemos deducir muchos resultados sobre $F_{ij}(t)$, en particular el tiempo promedio de primer arribo $\langle T_{ij} \rangle = \sum_{t=0}^{\infty} t F_{ij}(t)$. Si el reinicio se da al punto de partida ($r = i$), obtenemos:

$$\langle T_{ij} \rangle = \frac{\delta_{ij}}{P_j^\infty(i)} + \frac{1}{P_j^\infty(i)} \sum_{l=2}^N \frac{\langle j | \phi_l \rangle \langle \bar{\phi}_l | j \rangle - \langle i | \phi_l \rangle \langle \bar{\phi}_l | j \rangle}{1 - (1-\gamma)\lambda_l}. \quad (27)$$

²Ver Riascos *et. al* (2020) para detalles.

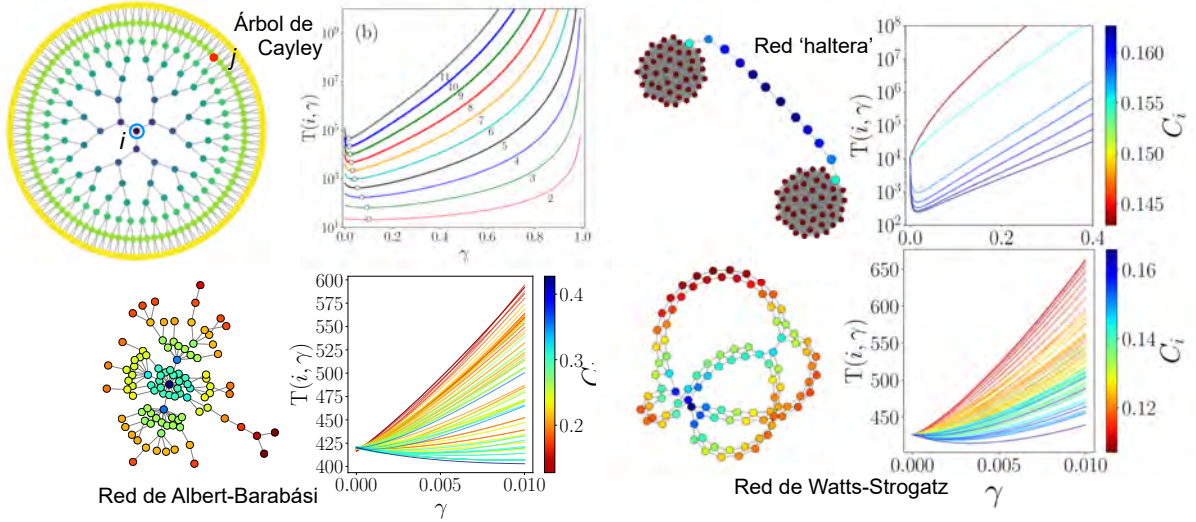


Figure 7: Caminatas con reinicio en varios tipos de redes [tomado de Riascos *et al.* 2020]. Excepto para el árbol de Cayley, $N = 100$ y el color de los nodos indica su centralidad: un nodo rojo se encuentra más lejos topológicamente de los otros nodos (centralidad C_i baja) que un nodo azul (C_i alto).

Los resultados (25) y (27) nos enseñan algo esencial: las propiedades de un proceso con reinicio se pueden deducir del conocimiento de la dinámica *sin* reinicio, caracterizada por la matriz \mathbf{W} . Si para una red dada se conoce $\{\lambda_l, |\phi_l\rangle, \langle \bar{\phi}_l|\}$, deducimos P_{ij} , $\langle T_{ij} \rangle$ o T_i muy rápidamente para cualquier valor del sitio de reinicio y probabilidad de reinicio γ . Podemos ahora aplicar estos resultados a varios ejemplos de redes, que se ilustran en la figura 7.

Los arboles de Cayley son redes donde los nodos tienen grado z , excepto los nodos más exteriores, de grado 1. En una red de Cayley de N nodos, la distancia topológica promedio entre dos nodos, definida como el número mínimo de pasos para ir de uno al otro, es $\propto \ln N$. Esta distancia es muy corta y caracteriza las redes llamadas de ‘mundo pequeño’ (contrariamente a una red $1d$, $2d$, etc...). El tiempo global de primer arribo T_i , donde i es el nodo central, es mínimo para un cierto valor de γ , que disminuye a medida que el número de capas aumenta. En un árbol infinito, si el blanco j se encuentra a una distancia d de i , $\langle T_{ij} \rangle$ es mínimo para $\gamma = \gamma^*$ con

$$\gamma^* \simeq \frac{1}{d} \left(\frac{z-2}{z} \right), \quad d \gg 1, \quad (28)$$

mientras que el tiempo mínimo correspondiente es

$$\langle T_{ij}^* \rangle \simeq \frac{N_d \ln N_d}{\ln(z-1)}, \quad (29)$$

donde N_d es el número de nodos a una distancia de i menor o igual a d , o $N_d = 1 + z \sum_{k=0}^{d-1} (z-1)^k \simeq \frac{z(z-1)^d}{z-2}$. La ecuación (29) nos dice que las búsquedas con reinicio óptimo son muy eficientes en árboles de Cayley, casi tan eficientes que la mejor estrategia posible, que consiste en visitar los nodos a una distancia d de manera sistemática. Se puede mostrar que esta última estrategia tiene un tiempo de búsqueda promedio $\simeq N_d$, sólo un poco menor que (29).

Las redes ‘halteras’ (ver figura 7) se componen de dos ‘comunidades’ de nodos conectados todos con todos, separadas por una cadena. En este caso, la posición del nodo de reinicio i tiene consecuencias dramáticas sobre la exploración de la red. Si i pertenece a una comunidad, el reinicio es muy desfavorable: T_i crece monótona y rápidamente con

γ . Si i es un sitio de la cadena, al contrario, T_i puede decrecer por ordenes de magnitud comparado con la caminata aleatoria $\gamma = 0$, antes de alcanzar un mínimo. Claramente, las búsquedas más rápidas son las que se reinician a nodos *centrales* de la red, es decir los que son cercanos (en promedio) a los otros nodos. Curiosamente, la centralidad del nodo inicial no juega ningún papel para la caminata aleatoria simple: cuando i varía, las diferentes curvas T_i tienen el mismo valor en $\gamma = 0$.

Unas conclusiones similares se pueden alcanzar con las redes de Watts-Strogatz, que son otro ejemplo de redes de mundo pequeño. Están formadas por un anillo unidimensional, con algunas conexiones que unen pares de nodos escogidas al azar ('atajos'). Los nodos cercanos a estas conexiones se vuelven centrales. Vemos cómo las variaciones de T_i con γ transitan de un comportamiento no-monótono a uno monótono creciente (y con valores siempre más grandes), conforme la centralidad del nodo de reinicio baja.

Contrariamente a los casos anteriores, las redes de Albert-Barabási son libres de escala: algunos pocos nodos tienen un gran número de conexiones, mientras que la mayoría tiene pocas. Las tendencias anteriores son menos claras en este ejemplo. En conclusión, estas exploraciones nos enseñan que el reinicio puede ser benéfico o no para la búsqueda, esto depende tanto de la posición del nodo de reinicio como de la arquitectura de la red.

5 Reiniciar partículas usando potenciales intermitentes

Una limitación de las teorías expuestas arriba es que asumen que el reinicio ocurre en un tiempo nulo o muy corto. Si tuvieramos que diseñar un experimento de reinicio con un sistema físico real, traer una partícula a un punto fijo tomaría algún tiempo. Además, la posición de reinicio no tendría que ser siempre la misma, podría fluctuar debido a restricciones físicas. De hecho, casi no se han realizado estudios experimentales sobre reinicio.

Aquí, veremos cómo prender y apagar un potencial confinante puede emular un proceso de reinicio sobre una partícula browniana microscópica. En la figura 8a, el potencial $v(x)$ de un sistema unidimensional se prende a una tasa r_0 y se apaga a tasa r_1 . En el estado apagado, la partícula se difunde libremente, mientras que en el prendido, siente un sesgo que la atrae hacia el origen (al mínimo del potencial) similarmente a un reinicio. Sin embargo, la partícula no se encuentra necesariamente en 0 cuando $v(x)$ se vuelve a apagar. Los intervalos de tiempo prendido/apagado se alternan y se distribuyen de manera exponencial, con promedios $1/r_1$ y $1/r_0$, respectivamente. En la posición $x = -1$, se encuentra un blanco que termina el proceso al primer encuentro con la partícula (ver trayectoria). ¿Cómo escoger las tasas (r_0, r_1) para minimizar este tiempo de primer arribo? Vemos en la figura 8b un ejemplo de minimización. ¿Cómo las tasas óptimas (r_0^*, r_1^*) dependen de $v(x)$ o de su dureza?

Este problema se podría estudiar experimentalmente con pinzas ópticas para crear un potencial definido, una técnica ampliamente utilizada para manipular partículas microscópicas. También se puede analizar teóricamente. Como son dos tasas en lugar de una, el problema es potencialmente más rico que el de Evans y Majumdar, pero bastante más complicado también... La ecuación de movimiento de la partícula browniana está dada por una ecuación de Langevin:

$$\frac{dX}{dt} = -\sigma(t)v'(X) + \xi(t), \quad (30)$$

que proviene de la ecuación de conservación del momento de Newton en presencia de la fuerza externa $-\sigma(t)v'(X)$, donde $\sigma(t) = 1$ si el potencial está prendido (y 0 si está

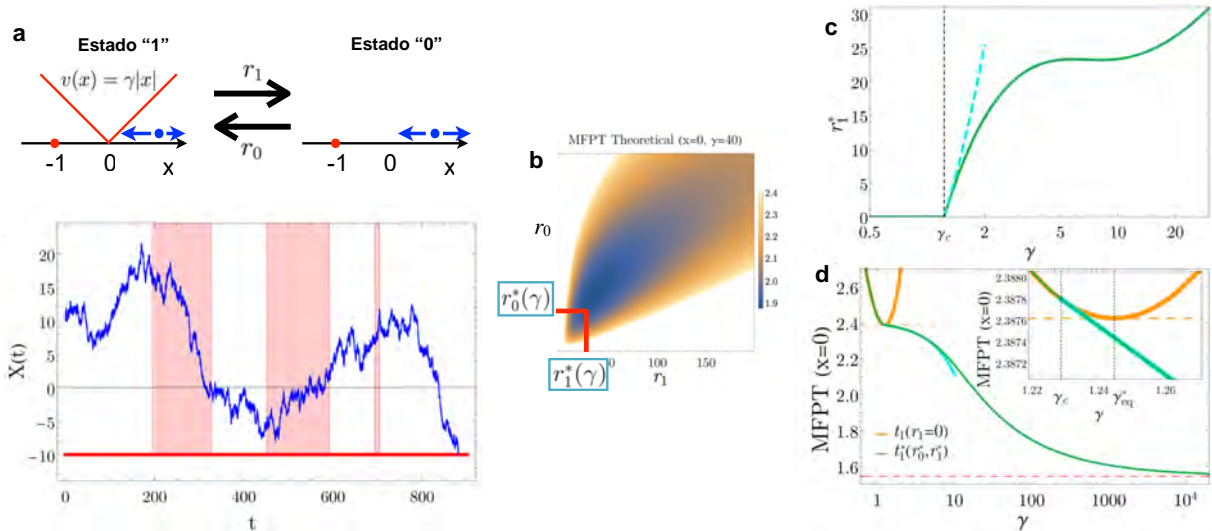


Figure 8: **a)** Una partícula sale de x a $t = 0$ y un potencial lineal confinante de dureza γ se prende y apaga estocásticamente, generando fases de difusión libre y sesgada (en rojo en la trayectoria) alternadas. El proceso termina en el primer encuentro con un blanco situado en -1 . **b)** Tiempo promedio de primer arribo (MFPT) a γ fijo como función de las tasas de prendido/apagado. Existe un mínimo en (r_0^*, r_1^*) . **c)** Tasa óptima r_1^* en función de γ . **d)** El MFPT correspondiente (curva verde). En naranja: MFPT con el potencial siempre prendido (tiempo de Kramers). Inset: zoom de la región de la transición de fase, cerca de $\gamma = \gamma_c$. [Algunas figuras tomadas de Mercado-Vásquez *et al.* (2020).]

apagado), de la fuerza de fricción de Stokes $-dX/dt$, y de una fuerza aleatoria o ruido blanco delta-correlacionado $\xi(t)$, donde $\langle \xi(t) \rangle = 0$ y $\langle \xi(t)\xi(t') \rangle = 2\delta(t - t')$. Esta fuerza aleatoria representa el efecto de las colisiones de las moléculas del fluido con la partícula browniana. En la ecuación de Langevin, se desprecia como es costumbre el término de aceleración $m\ddot{X}$. Nótese que aquí todas las variables y constantes son adimensionales. En ausencia de potencial, la partícula sigue un movimiento browniano. Recordamos que se trata de una trayectoria continua pero no derivable en ningún punto.

En (30), $\sigma(t)$ es un ruido binario ‘externo’, creado por el manipulador y gobernado por las tasas de transición r_0 y r_1 . Dada una posición inicial x , existen dos estados iniciales posibles para el potencial. Definimos $Q_1(x, t)$ como la probabilidad de sobrevivencia al tiempo t de la partícula empezando en x y con el potencial prendido a $t = 0$, mientras $Q_0(x, t)$ representa la misma cantidad con el potencial inicialmente apagado. Se puede mostrar que Q_0 y Q_1 satisfacen dos ecuaciones ‘backward’ de Fokker-Planck acopladas:

$$\frac{\partial Q_0}{\partial t} = \frac{\partial^2 Q_0}{\partial x^2} + r_0(Q_1 - Q_0), \quad (31)$$

$$\frac{\partial Q_1}{\partial t} = \frac{\partial^2 Q_1}{\partial x^2} - v'(x)\frac{\partial Q_1}{\partial x} + r_1(Q_0 - Q_1). \quad (32)$$

Estas ecuaciones no parecen tener una solución exacta en el caso general. Sin embargo, para el potencial lineal confinante, $v(x) = \gamma|x|$ con $\gamma > 0$, sí se pueden resolver parcialmente. Por lo menos, se pueden calcular los tiempos de primer arribo asociados $\langle t_0 \rangle = \int_0^\infty Q_0(x, t)dt$ y $\langle t_1 \rangle = \int_0^\infty Q_1(x, t)dt$ de manera exacta en función de x y de los parámetros γ , r_0 y r_1 . Aún en este caso, las expresiones son muy largas y complicadas, aunque involucren sólo funciones elementales. Queda fuera del contexto de este texto presentarlas y nos limitaremos a resumir algunos resultados entre los muchos que se pueden obtener.

Supongamos que $x = 0$ (la partícula sale del mínimo del potencial a $t = 0$) y con el potencial inicialmente prendido. Si la dureza del potencial γ es inferior a un valor crítico

altamente no-trivial, que se puede calcular exactamente como $\gamma_c = 1.228780\dots$, el valor óptimo de r_1 es simplemente $r_1^* = 0$. Es decir, la mejor estrategia de búsqueda consiste en no apagar el potencial nunca. Identificamos el tiempo de búsqueda asociado con el bien conocido tiempo de Kramers $\langle t_1^{(K)} \rangle$ requerido para brincar por encima de una barrera de potencial de altura γ en el equilibrio termodinámico (o para potenciales estacionarios). Obtenemos

$$\langle t_1^* \rangle(\gamma) = \langle t_1^{(K)} \rangle = \frac{2}{\gamma^2}(e^\gamma - 1) - \frac{1}{\gamma}, \quad \gamma < \gamma_c. \quad (33)$$

Recuperamos la ley de Arrhenius conocida para tasas de reacciones, o $1/\langle t_1^{(K)} \rangle \propto e^{-\gamma}$ cuando $\gamma \gg 1$. No es de sorprender, dado que las reacciones químicas son esencialmente procesos de primer arribo con barreras de activación. La forma analítica del tiempo $\langle t_1^{(K)} \rangle$ se muestra en la figura 8d en color naranja. Es válida para cualquier valor de γ y es no-monótona.

Si la dureza del potencial excede γ_c , entonces existe un protocolo de reinicio no-trivial, con $r_1^* > 0$ y $r_0^* > 0$, que minimiza el tiempo promedio de búsqueda. Las variaciones de r_1^* con γ se muestran en la figura 8c: son típicas de una transición de fase continua de segundo orden. Para γ arriba pero cerca de γ_c , la tasa óptima de apagado crece linealmente,

$$r_1^*(\gamma) \simeq A(\gamma - \gamma_c)^\beta \quad \text{con } \beta = 1 \text{ y } A = 32.913015\dots \quad (34)$$

A partir de γ_c , el tiempo mínimo de búsqueda $\langle t_1^* \rangle(\gamma)$ se vuelve *inferior* al tiempo de Kramers, ver figura 8d. El hecho de sacar el sistema de equilibrio (prendiendo y apagando continuamente el potencial) permite entonces mejorar la eficiencia de la búsqueda. Mientras que $\langle t_1^{(K)} \rangle$ crece exponencialmente a γ grande, $\langle t_1^* \rangle(\gamma)$ disminuye monótonamente, hasta alcanzar el valor asintótico 1.5441... que es bastante más bajo incluso que $\min_\gamma \langle t_1^{(K)} \rangle$. Cuando $\gamma \rightarrow \infty$, el potencial es tan fuerte que trae de manera instantánea la partícula de regreso al origen y recuperamos el problema de Evans y Majumdar. Enviamos al lector a Mercado-Vasquez *et al.* (2020) para mayores detalles.

6 Conclusiones

Hicimos una breve revisión de los procesos de reinicio y de algunas de sus aplicaciones. En combinación con la difusión, el reinicio constituye una estrategia eficiente en muchos problemas de búsqueda aleatoria. Además, se generan efectos estáticos y dinámicos ricos, que no se observan en el equilibrio termodinámico. Todavía queda mucho por explorar, tanto en el ámbito teórico como experimental.

Referencias

Libros sobre teoría de probabilidad y difusión:

C. W. Gardiner, *Handbook of Stochastic Methods: For Physics, Chemistry and the Natural Sciences* (Springer, New York, 1996).

L. E. Reichl, *A Modern Course in Statistical Physics* (Wiley, New York, 1998).

H. Risken, *The Fokker-Planck Equation* (Springer, Berlin, 1984).

Sobre teoría de procesos de primer arribo:

S. Redner, *A Guide to First-Passage Processes* (Cambridge University Press, Cambridge, 2001).

H. Risken, *The Fokker-Planck Equation* (Springer, Berlin, 1984).

Búsquedas aleatorias:

G. M. Viswanathan, M. G. E. da Luz, E. P. Raposo and H. E. Stanley, *The Physics of Foraging: An Introduction to Random Searches and Biological Encounters* (Cambridge University Press, 2011).

O. Bénichou, C. Loverdo, M. Moreau, and R. Voituriez, Intermittent search strategies, *Rev. Mod. Phys.* **83**, 81–129 (2011).

Procesos de reinicio:

M. R. Evans and S. N. Majumdar, Diffusion with stochastic resetting, *Phys. Rev. Lett.* **106**, 160601 (2011).

M. R. Evans, S. N. Majumdar and G. Schehr, Stochastic resetting and applications, *J. Phys. A: Math. Theor.* **53**, 193001 (2020) [artículo de revisión].

B. Besga, A. Bovon, A. Petrosyan, S. N. Majumdar and S. Ciliberto, Optimal mean first-passage time for a Brownian searcher subjected to resetting: Experimental and theoretical results, *Phys. Rev. Research* **2**, 032029(R) (2020).

Seminario en pdf: https://zakopane.if.uj.edu.pl/event/9/contributions/319/attachments/42/23/Satya_N._Majumdar.pdf

El uso de potenciales intermitentes para reiniciar partículas browniana:

G. Mercado-Vásquez, D. Boyer, S. N. Majumdar and G. Schehr, Intermittent resetting potentials, *J. Stat. Mech.* 113203 (2020).

Referencias generales sobre redes complejas:

M. E. J. Newman, The structure and function of complex networks, *SIAM Rev.* **45**, 167–256 (2003).

A.-L. Barabási, *Network Science*, 2016 (Creative Commons license) ISBN: 9781107076266.

Artículo de revisión sobre caminatas aleatorias en redes:

N. Masuda, M. A. Porter and R. Lambiotte, Random walks and diffusion on networks, *Phys. Rep.* 716–717, 1 (2017).

Caminatas aleatorias en redes con reinicio estocástico:

A. P. Riascos, D. Boyer, P. Herringer and J. L. Mateos, Random walks on networks with stochastic resetting, *Phys. Rev. E* **101**, 062147 (2020).



Medición precisa de la polarización de la luz

Neil C. Bruce

*Instituto de Ciencias Aplicadas y Tecnología,
Universidad Nacional Autónoma de México,
Ciudad Universitaria,
Ciudad de México.*

Introducción

La luz es una onda vectorial electromagnética transversal. Los campos eléctricos y magnéticos son perpendiculares a la dirección de propagación, con una amplitud, una fase y una dirección en cada punto del espacio [1]. Las ecuaciones de Maxwell relacionan los campos eléctricos y magnéticos; sabiendo la dirección y amplitud del campo eléctrico, sabemos la amplitud y dirección del campo magnético, entonces sabiendo el campo eléctrico, conocemos el comportamiento de toda la onda electromagnética. Al comportamiento de la dirección y amplitud del campo eléctrico con la propagación de la onda se le conoce como **polarización** de la luz [1].

Para describir la polarización, se descompone el campo eléctrico en dos componentes perpendiculares [1]. En esta descripción hay tres parámetros: la amplitud de cada componente y la fase relativa entre los componentes. Para incluir el caso de luz no polarizada, se requiere también incluir la intensidad total del haz, lo que significa que, para describir cualquier estado de polarización necesitamos cuatro parámetros. Esto también significa que, para medir completamente la polarización de un haz, tenemos que medir estos mismos cuatro parámetros. Hay varias formas matemáticas para escribir estos cuatro parámetros, pero los más comunes son el formalismo de Jones [2,3] y los parámetros de Stokes [4].

El vector de Jones utiliza la amplitud y fase de cada componente del campo eléctrico perpendicular para describir la polarización:

$$J = \begin{pmatrix} E_x e^{i\phi_x} \\ E_y e^{i\phi_y} \end{pmatrix}, \quad (1)$$

y la acción de un componente óptico cambia el vector de Jones de un valor inicial J_1 a un valor final J_2 , a través de una matriz de 2×2 , llamado la matriz de Jones:

$$J_2 = \begin{pmatrix} m_{11} & m_{12} \\ m_{21} & m_{22} \end{pmatrix} J_1. \quad (2)$$

El vector de Jones describe la polarización de un haz, y la matriz de Jones describe la acción de un componente óptico o una muestra sobre un estado de polarización, y requieren el conocimiento de la amplitud y fase de los componentes perpendiculares del campo eléctrico, entonces no son fácilmente medibles. También, no pueden describir la luz no-polarizada, luz sin una polarización preferencial.

El formalismo del vector de Stokes es un método que trabaja con elementos medibles de intensidad. El vector de Stokes describe la polarización de un haz, y su definición es [1]:

$$S = \begin{pmatrix} S_0 \\ S_1 \\ S_2 \\ S_3 \end{pmatrix} = \begin{pmatrix} I_x + I_y \\ I_x - I_y \\ I_{+45} + I_{-45} \\ I_{CD} + I_{CI} \end{pmatrix}, \quad (3)$$

donde I_x es la intensidad en el componente de polarización en la dirección x , I_y es la intensidad en el componente de polarización en la dirección y , I_{+45} es la intensidad en el componente de polarización a 45° con respecto a la dirección x , I_{-45} es la intensidad en el componente de polarización a -45° con respecto a la dirección x , y I_{CD} y I_{CI} son los componentes de polarización circular derecha e izquierda, respectivamente. Para separar estos componentes se utilizan polarizadores, que dejan pasar el campo eléctrico en una sola dirección, y retardadores, que cambian la fase entre los dos componentes perpendiculares.

El efecto de un componente óptico o una muestra sobre la polarización, que cambia un vector de Stokes a otro vector de Stokes, es una matriz 4×4 llamado la matriz de Mueller:

$$S_2 = \begin{pmatrix} M_{11} & M_{12} & M_{13} & M_{14} \\ M_{21} & M_{22} & M_{23} & M_{24} \\ M_{31} & M_{32} & M_{33} & M_{34} \\ M_{41} & M_{42} & M_{43} & M_{44} \end{pmatrix} S_1. \quad (4)$$

La medición del vector de Stokes o la matriz de Mueller se llama polarimetría. Las aplicaciones de polarimetría son muy amplias. Por ejemplo, se ha utilizado para caracterizar imágenes médicas [5-9], en percepción remota [10,11], en la caracterización de materiales [12,13], y astronomía [14-16], entre otros.

En este trabajo presentamos un análisis numérico de un polarímetro de Mueller y su sensibilidad a errores experimentales y ruido de medición.

Polarimetría de Stokes

La polarimetría de Stokes mide los componentes de estados de polarización definidos para poder determinar la polarización desconocido en un haz. Hay 4 parámetros de Stokes desconocidos, entonces un mínimo de 4 mediciones es requerido para fijar la polarización desconocido. Hay diferentes técnicas experimentales para los polarímetros [1], como modulación de polarización o de polarimetría de Fourier, división de amplitud, división de abertura, o la técnica que estamos utilizando en el ICAT: secuencia temporal. La técnica de secuencia temporal para un polarímetro de Stokes utiliza al menos 4 configuraciones del polarímetro, una después de la otra, para obtener las 4 mediciones requeridas. En nuestro caso estamos usando retardadores variables de cristales líquidos (LCVRs por sus siglas en inglés “Liquid-Crystal Variable Retarder”), que tiene una variación de la retardancia con el voltaje aplicado. El arreglo experimental se muestra en la Figura 1.

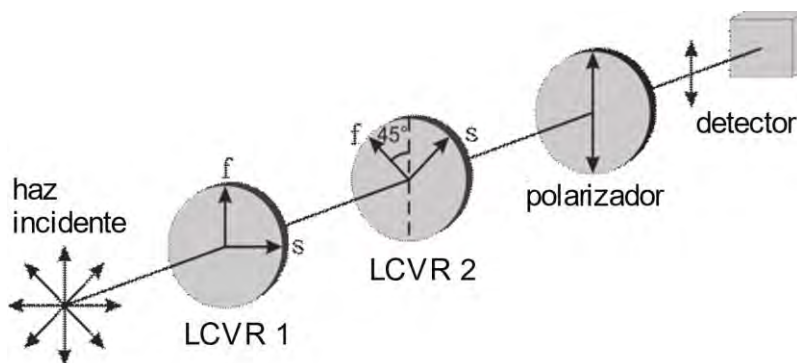


Figura 1: Arreglo experimental para un polarímetro de Stokes

El polarímetro utiliza dos LCVRs, con sus ejes rápidos a 0° y 45°, respectivamente, y un polarizador con su eje de transmisión en 0°. Cambiando el voltaje a los dos retardadores, se puede cambiar la polarización detectada.

Escribiendo el vector de Stokes del haz incidente como S_{inc} , el vector de Stokes del haz que llega al detector se escribe como:

$$S_{1det} = M_{pol0} M_{R12,45} M_{R11,0} S_{inc}, \quad (5)$$

donde M_{pol0} es la matriz de Mueller para un polarizador a 0° , y $M_{Rij,\theta}$ es la matriz de Mueller para el LCVR j , con retardancia δ_{ij} , y con su eje rápida a un ángulo θ . Para la primera configuración de retardancias, y en términos de los componentes de la matriz de Mueller total del sistema y los vectores de Stokes:

$$S_{1det} = \begin{pmatrix} S_{1det0} \\ S_{1det1} \\ S_{1det2} \\ S_{1det3} \end{pmatrix} = \begin{pmatrix} M_{111} & M_{112} & M_{113} & M_{114} \\ M_{121} & M_{122} & M_{123} & M_{124} \\ M_{131} & M_{132} & M_{133} & M_{134} \\ M_{141} & M_{142} & M_{143} & M_{144} \end{pmatrix} S_{inc}. \quad (6)$$

De la ecuación (3), la energía total en un haz es igual al primer término del vector de Stokes, entonces la energía detectada es igual a S_{det0} , entonces:

$$S_{1det0} = I_1 = (M_{111} \quad M_{112} \quad M_{113} \quad M_{114}) S_{inc}. \quad (7)$$

Ahora, necesitamos usar al menos 4 configuraciones de retardancias, para tener 4 mediciones de intensidad:

$$I = \begin{pmatrix} I_1 \\ I_2 \\ I_3 \\ I_4 \end{pmatrix} = \begin{pmatrix} M_{111} & M_{112} & M_{113} & M_{114} \\ M_{211} & M_{212} & M_{213} & M_{214} \\ M_{311} & M_{312} & M_{313} & M_{314} \\ M_{411} & M_{412} & M_{413} & M_{414} \end{pmatrix} S_{inc} = M S_{inc}, \quad (8)$$

y se puede recuperar el vector de Stokes del haz incidente utilizando el inverso de la matriz M :

$$S_{inc} = M^{-1} I \quad (9)$$

La matriz M se llama la matriz característica del polarímetro.

El número de condición (CN por sus siglas en inglés, "condition number") es un parámetro que indica que tan sensible es el inverso a cambios en los valores de los elementos de una matriz [17]. El CN se calcula como la razón entre el valor singular más grande de la matriz y el valor singular más pequeño. Un CN mayor indica que el inverso tendrá más error con errores en los valores de la matriz original. En el caso de polarimetría, los errores en los elementos de la matriz M corresponden a errores experimentales en el instrumento, tal como errores en la alineación de los ejes de los componentes o errores en las retardancias aplicadas.

Tyo [18] mostró que el valor óptimo del CN para medir los 4 parámetros del vector de Stokes es $\sqrt{3} = 1.732$. Algunos ejemplos de matrices características y sus CN son: primero midiendo 4 intensidades con polarizaciones lineal horizontal, lineal vertical, lineal a $+45^\circ$, y circular derecha:

$$M = \begin{pmatrix} 1 & 1 & 0 & 0 \\ 1 & -1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 \end{pmatrix}, \quad \text{CN} = 3.2255, \quad (10)$$

que claramente no es optimizada. Un sistema con 6 mediciones, usando las mediciones del caso anterior y agregando polarización lineal a -45° y polarización circular izquierda da:

$$M = \begin{pmatrix} 1 & 1 & 0 & 0 \\ 1 & -1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & -1 & 0 \\ 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & -1 \end{pmatrix}, \quad \text{CN} = 1.732, \quad (11)$$

que sí es optimizada. La diferencia entre estos dos casos es que el segundo caso, ecuación (11), mide todas las intensidades requeridas en la ecuación (3) independientemente, mientras en la ecuación (10) no se mide todas las intensidades, y hay que usar algunos de los elementos del vector de Stokes para calcular los otros elementos, y errores propagan de una forma mayor. Sin embargo, sí hay combinaciones de 4 mediciones que dan una matriz característica optimizada. Por ejemplo [19]:

$$M = \begin{pmatrix} 1 & -0.9426 & -0.3338 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0.4278 & -0.3322 & 0.8162 \\ 1 & 0.4716 & -0.3338 & -0.8126 \end{pmatrix}, \quad \text{CN} = 1.732. \quad (12)$$

Presentaremos los parámetros del polarímetro para obtener esta matriz característica más adelante.

Polarimetría de Mueller

Para medir una matriz de Mueller, que tiene 16 elementos desconocidos, se requiere un mínimo de 16 mediciones. Se logra esto con una combinación de 4 estados de polarización incidentes sobre la muestra y la medición de 4 estados de polarización. Esto es, en este caso debemos controlar la polarización incidente y la detectada. La detección es con el mismo sistema del polarímetro de Stokes y se llama un analizador del estado de polarización (PSA por sus siglas en inglés

“Polarization State Analyzer”), y la polarización incidente se genera con un PSA invertido, como se muestra en la Figura 2, y se llama un generador de estados de polarización (PSG por sus siglas en inglés “Polarization State Generator”).

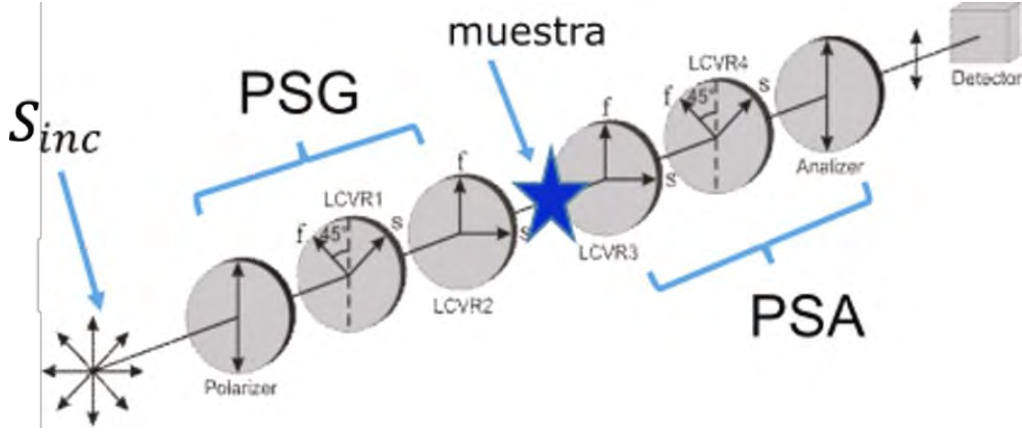


Figura 2: Arreglo experimental para un polarímetro de Mueller

La matriz de Mueller que describe el polarímetro de Mueller es:

$$M_M = M^{PSA} M^{Muestra} M^{PSG}, \quad (13)$$

y el vector de Stokes que llega al detector para cada configuración del polarímetro es:

$$S_{det} = \begin{pmatrix} S_{det0} \\ S_{det1} \\ S_{det2} \\ S_{det3} \end{pmatrix} = M_M S_{inc}, \quad (14)$$

y, otra vez, la intensidad medida para cada configuración es S_{det0} . Suponiendo luz no-polarizado incidente y escribiendo S_{det0} en términos de los componentes de las matrices, se puede escribir:

$$S_{det0} = \sum_{i=1}^4 \sum_{j=1}^4 M_{1i}^{PSA} M_{j1}^{PSG} M_{ij}^{Muestra}, \quad (15)$$

que se puede escribir como:

$$S_{det0}^1 = W^1 M'^{Muestra}, \quad (16)$$

donde el superíndice 1 indica que es para la primera configuración del polarímetro, y:

$$W^1 = (M_{11}^{PSA} M_{11}^{PSG} \quad M_{11}^{PSA} M_{21}^{PSG} \quad M_{11}^{PSA} M_{31}^{PSG} \quad M_{11}^{PSA} M_{41}^{PSG} \quad M_{12}^{PSA} M_{11}^{PSG} \\ M_{12}^{PSA} M_{11}^{PSG} \quad \dots \quad M_{14}^{PSA} M_{11}^{PSG} \quad M_{14}^{PSA} M_{21}^{PSG} \quad M_{14}^{PSA} M_{31}^{PSG} \quad M_{14}^{PSA} M_{41}^{PSG}), \quad (17)$$

$$M'^{Muestra} = \\ (M_{11}^{Muestra} \quad M_{12}^{Muestra} \quad M_{13}^{Muestra} \quad M_{14}^{Muestra} \quad M_{21}^{Muestra} \quad \dots \quad M_{44}^{Muestra})^T \quad (18)$$

donde el superíndice T indica la transpuesta del vector. Ahora, escribiendo las ecuaciones para las 16 configuraciones que se requiere para medir la matriz de Mueller desconocida de:

$$S = \begin{pmatrix} S_{det0}^1 \\ S_{det0}^2 \\ S_{det0}^3 \\ \vdots \\ S_{det0}^{16} \end{pmatrix} = \begin{pmatrix} W^1 \\ W^2 \\ W^3 \\ \vdots \\ W^4 \end{pmatrix} M'^{Muestra} = W M'^{Muestra}, \quad (19)$$

y el vector con los elementos de la matriz de Mueller desconocida es:

$$M'^{Muestra} = W^{-1} S. \quad (20)$$

Usualmente, la optimización de un polarímetro de Mueller se hace optimizando el PSG y el PSA por separado, ya que es más sencillo que optimizar la matriz W .

Calibración de las mediciones

Debido a los errores experimentales que siempre están presentes, se requiere un método de calibración. En este trabajo se utiliza el método de eigenvalores desarrollado por Compain [20], que es uno de los más populares en la literatura. Este método es un procedimiento de calibración general para calibrar polarímetros utilizando varias muestras de referencia conocidas. Utiliza un mapeo de una muestra ideal y las intensidades medidas, y encuentra la mejor solución para la matriz característica a través de un análisis de álgebra lineal (eigenvalores). Los errores en el polarímetro y en las muestras de calibración son calculados en el análisis.

Simulación

La simulación de un polarímetro de Mueller usando LCVR consiste en los siguientes pasos:

1. Calcular la matriz W usando la configuración optimizada con los ángulos de los ejes rápidos de LCVR1 y LCVR4 en 27.4° , los ejes rápidos de LCVR2 y LCVR3 en 72.4° , y las retardancias en los LCVRs igual a 135° y 315° . Todas las combinaciones de las 2 retardancias en los 4 LCVR da las 16 mediciones requeridas. Estos son los valores usados para obtener la matriz característica de la ecuación (12).
2. Usando una muestra conocida se calculan las 16 intensidades medidas para las 16 configuraciones del polarímetro, pero incluyendo errores experimentales en las retardancias y posiciones de los ejes de los LCVR, y los polarizadores. Además, se agrega ruido aleatorio Gaussiano a las intensidades, para simular un experimento real.
3. Usando las 16 intensidades calculados en el paso 2, y la matriz W ideal, calculado en el paso 1, se calcula la matriz de Mueller de la muestra con la ecuación (20). Este paso da una estimación de la matriz de Mueller no-calibrada.
4. Se aplica el método de calibración de eigenvalores [20] a los resultados obtenidos, utilizando las intensidades calculadas para los cuatro casos: aire (sin muestra), un polarizador lineal horizontal, un polarizador lineal vertical, y un retardador de cuarto de onda a 45° . Este paso da una nueva estimación de la matriz de Mueller calibrada.
5. Se compara las dos estimaciones de la matriz de Mueller (calibrada y no-calibrada) con la conocida introducida en la simulación. Dado que utilizamos muestras conocidas rotando, en este paso se calcula el error total rms dado por:

$$rms = \sqrt{\sum_{\substack{\text{todas las} \\ \text{muestras}}} \sum_{\substack{\text{ángulo de} \\ \text{rotación}}} \sum_{\substack{\text{de parametro} \\ \text{de Stokes}}} \frac{1}{4N} (S^{est} - S^{ideal})^2} \quad (21)$$

6. Se repite desde el paso 2, con diferentes errores aleatorios.

Resultados

La Figura 3 muestra resultados típicos, para un nivel de ruido Gaussiano de 1% de la señal, antes de la calibración y después de la calibración. Se gráfica el CN del polarímetro contra el error total rms, y cada punto en las gráficas representa el polarímetro con unos errores experimentales específicos. El CN es el número de

condición del polarímetro con los errores experimentales introducidos en la simulación, y el polarímetro ideal, sin errores tiene $CN = 1.732$. Así, el CN es una medida de la magnitud de todos los errores experimentales en el polarímetro. Nota que diferentes valores de los errores pueden dar el mismo CN, entonces las gráficas tienen múltiples puntos para cada CN.

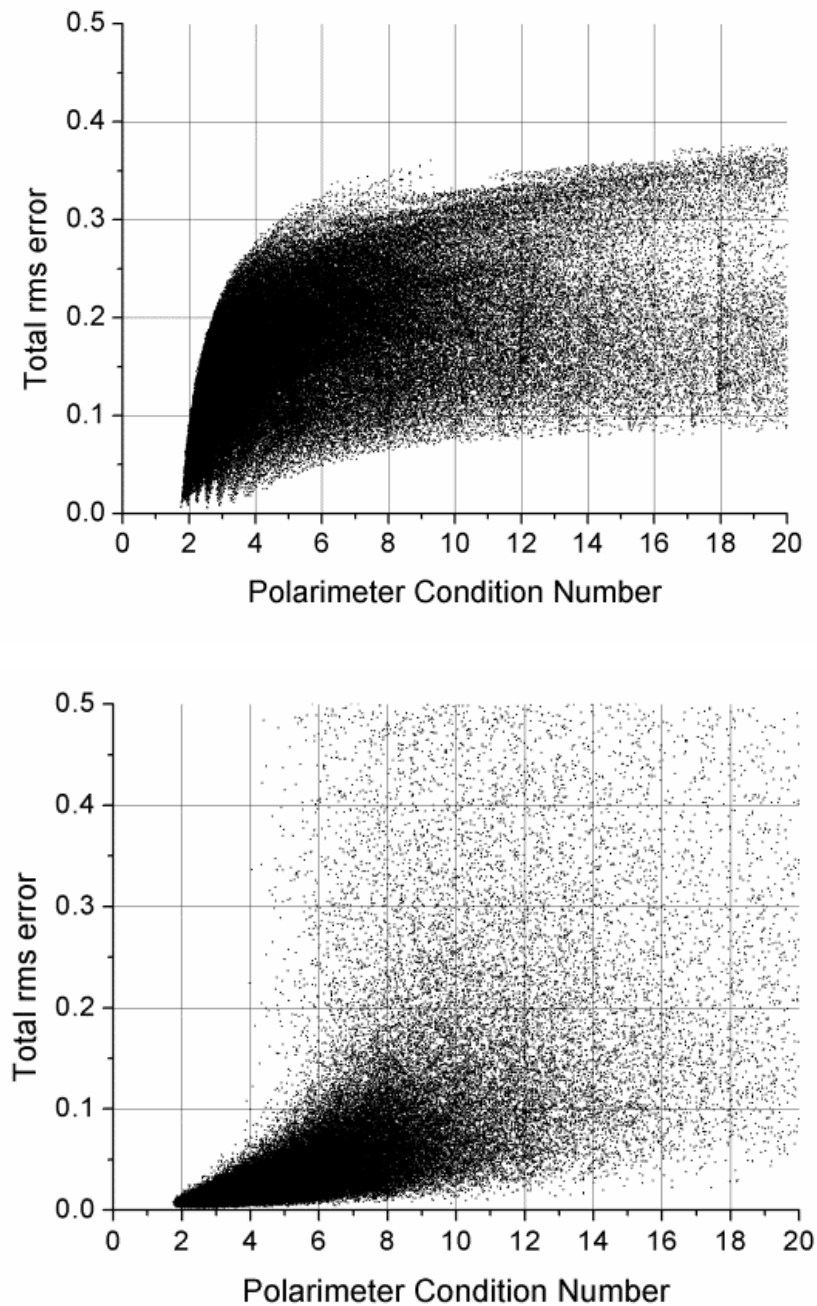


Figura 3: el error total rms contra el CN del polarímetro con errores experimentales para: las matrices de Mueller recuperadas no-calibradas (arriba), y calibradas (abajo).

De la Figura 3 (arriba) se puede ver que [21], antes de la calibración, el error total rms tiene un rango amplio de valores para todos los valores de CN, aunque para valores de CN cerca de optimizados (1.732) el rango de error total rms es menor. También, antes de la calibración el error total rms tiende a aumentar ligeramente con un aumento en el CN (un aumento en los errores experimentales). Después de la calibración de eigenvalores, Figura 3 (abajo), cambia la distribución de los puntos. Para valores bajas de CN, debajo de $CN \approx 5$, el error total rms es reducido en todos los casos, mostrando que la calibración está corrigiendo los errores introducidos en las mediciones por los errores experimentales. Para $CN > 5$ la situación cambia. En este caso el método de calibración corrige algunos de los casos (hay más puntos en valores más bajos del error total rms), pero empeora otros casos (más puntos con valores del error total rms). Esto significa que el éxito del método de calibración de eigenvalores depende de la configuración de los errores experimentales en el polarímetro para $CN > 5$.

La Figura 4 muestra la corrección de las matrices de Mueller, para diferentes valores de ruido de medición, a través del parámetro:

$$Drms = rms_{antes} - rms_{después}$$

donde los subíndices "antes" y "después" refiere a antes y después del proceso de calibración. Entonces si $rms_{antes} = rms_{después}$, no hay mejora en los resultados con el método de calibración, y entonces $Drms = 0$; y si $rms_{después} = 0$, esto es, cuando la calibración corrige completamente el error, $Drms = rms_{antes}$. Este segundo caso se muestra con la línea gris inclinada en las gráficas de la Figura 4. También, en todas las gráficas de la Figura 4 se puede ver que hay muchos puntos con un valor de $Drms < 0$, esto significa que $rms_{después} > rms_{antes}$, y otra vez, el proceso de calibración empeora los resultados de las matrices de Mueller reconstruidas. Los valores de rms_{antes} máximo que muestra ningún valor negativo $Drms$ se muestran con las líneas gris verticales en las gráficas de la Figura 4. Para el caso de 5% de ruido Gaussiano, no hay ningún valor de rms_{antes} que no muestra valores negativos de $Drms$, entonces en este caso no hay ningún nivel de errores experimentales que garanticen que el proceso de calibración funciona correctamente. Por esta razón, el caso de polarímetros para niveles de luz bajos todavía es un área de estudio importante.

También en la Figura 4, con la escala y las curvas en rojo, se muestran el porcentaje de casos en el cual el algoritmo de la calibración no funciona. Esto incluye los casos que dan $Drms < 0$ y los casos en los cuales el algoritmo no termina. Se puede ver que en todos los casos el porcentaje de falla del algoritmo aumenta con un aumento en rms_{antes} . Estos resultados muestran que hay límites para la aplicación del método de calibración.

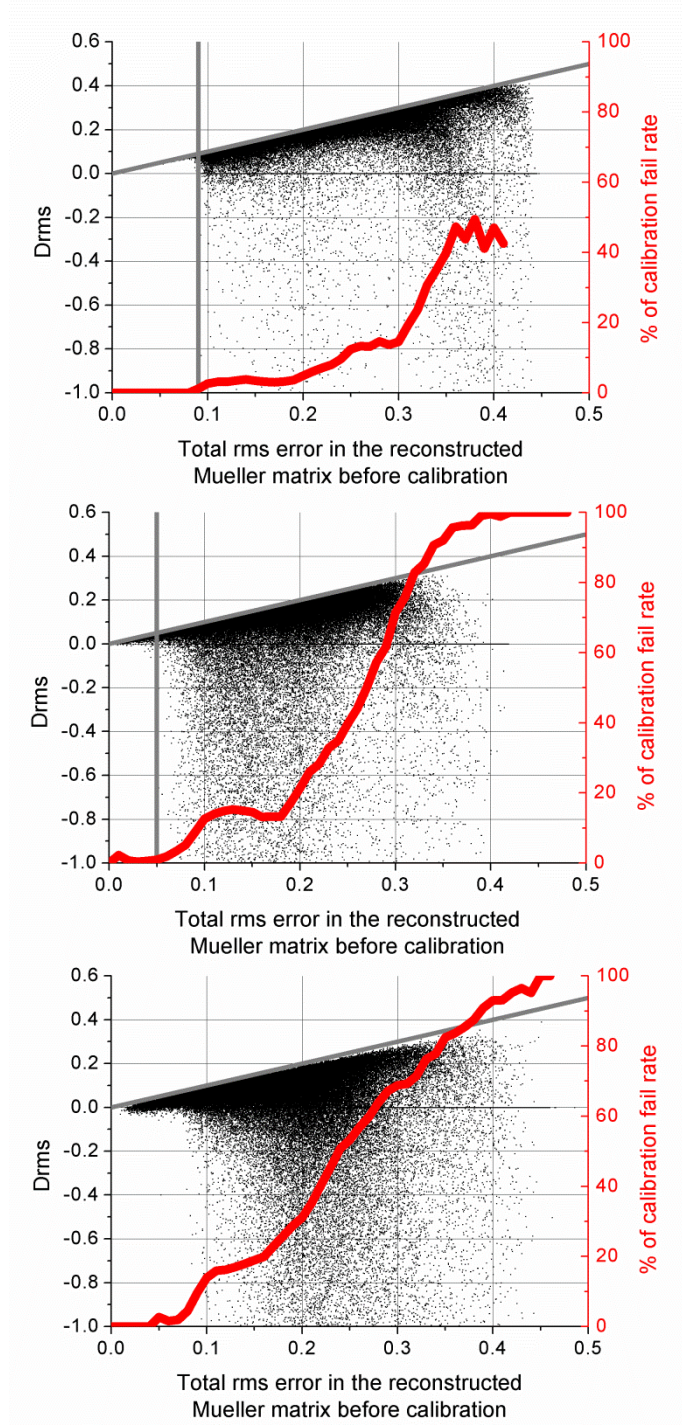


Figura 4: gráficas de $Drms$ vs rms_{antes} , con los puntos negros, para 0.5% de ruido de medición (arriba), 1.5% de ruido (centro), y 5% de ruido (abajo). También se muestra, en rojo, el porcentaje de falla del método de calibración vs rms_{antes} para los mismos casos.

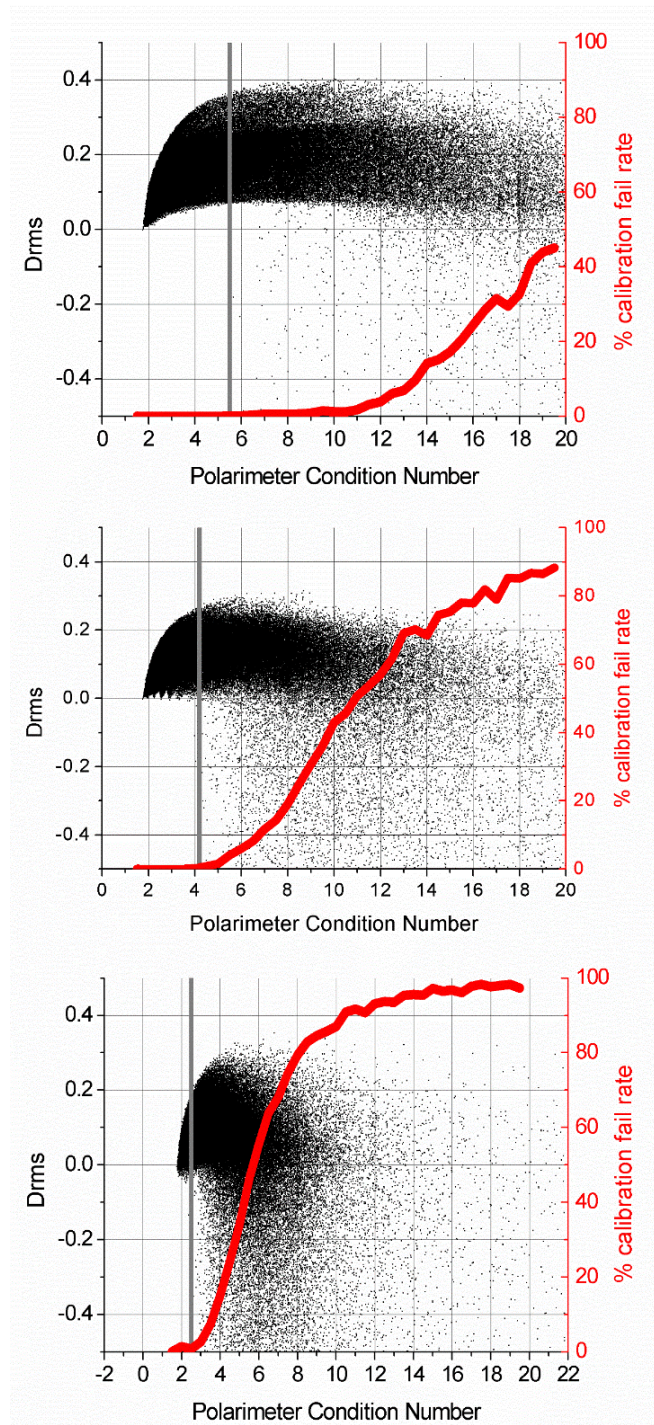


Figura 5: los mismos resultados que la Figura 4, pero graficados contra el CN del polarímetro con errores experimentales, para mostrar los valores límites de CN, para los cuales el método de calibración funciona correctamente.

La Figura 5 muestra los mismos resultados que la Figura 4, pero como función del CN del polarímetro con errores experimentales [21], y estos resultados se resumen en la Figura 6, que muestra el valor límite de CN para el cual el proceso de calibración funciona, como función del nivel de ruido Gaussiano de las mediciones. Como se esperaba, cuando hay más ruido en las mediciones, el sistema experimental debe estar más cerca a optimizado para que pueda funcionar correctamente el proceso de calibración de eigenvalores. La Figura 7 muestra la relación entre el valor del CN y los valores promedios de los errores experimentales para retardancias y los ángulos de los ejes rápidos de los LCVRs, para relacionar los CN límites de la Figura 6 con los errores experimentales. Se puede ver de la Figura 7 que los errores en los ángulos de los ejes rápidos son más importantes, ya que el CN aumenta más rápido con este error que con el error en las retardancias.

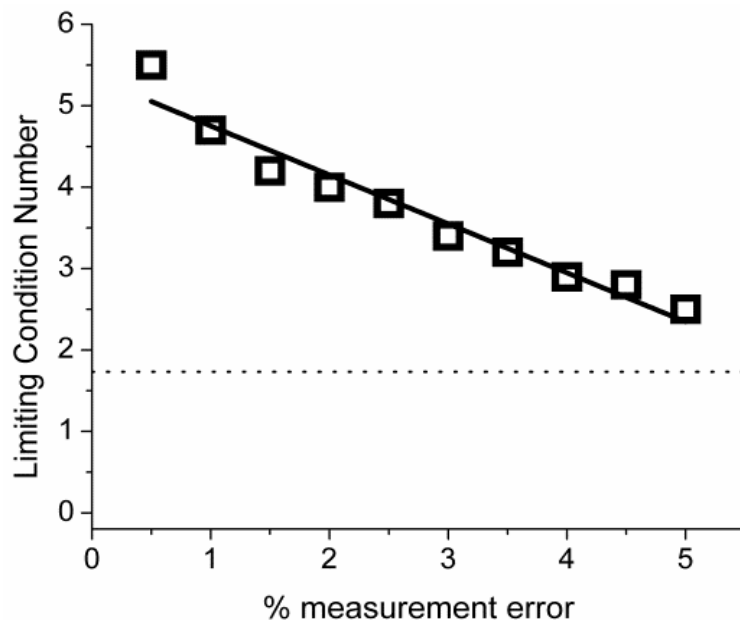


Figura 6: límite del CN para el cual funciona correctamente el proceso de calibración, como función del nivel de ruido Gaussiano de las mediciones.

También es importante calcular los errores experimentales permitidos, esto es la tolerancia en los errores experimentales, para un error total rms requerido en las matrices de Mueller recuperadas. La Figura 8 muestra dos ejemplos de los límites en el CN para errores totales rms iguales a 0.1, 0.05 y 0.02, para diferentes niveles de ruido de medición [22]. Los valores del CN límite se reducen como aumenta el ruido de medición, y como se reduce la presión requerido en las matrices de Mueller recuperadas, como se muestra en la Figura 9. Otra vez las tolerancias sobre los ángulos de los ejes rápidos de los LCVRs son más estrictos que las tolerancias sobre las retardancias.

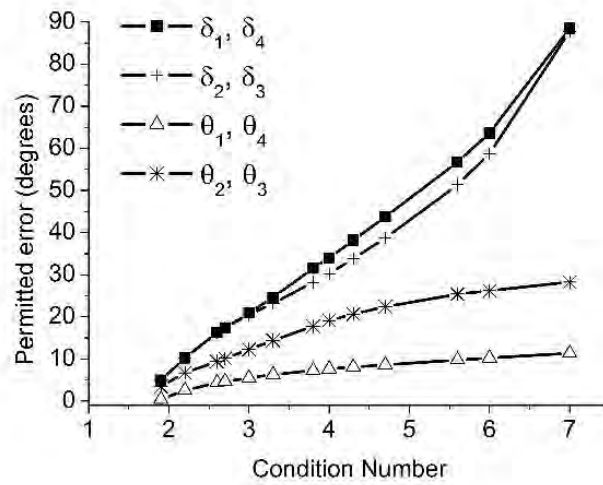


Figura 7: Relación entre el CN y los valores promedio de los errores experimentales de los LCVRs.

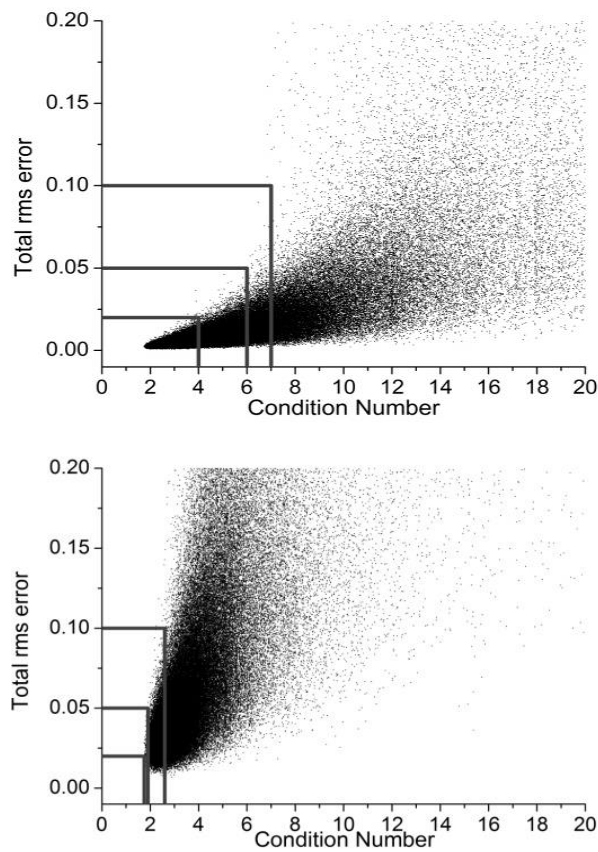


Figura 8: gráficas del error total rms contra el CN, mostrando el límite en el CN (líneas grises verticales), para valores límites del error total rms (líneas grises horizontales) de 0.1, 0.05 y 0.02.

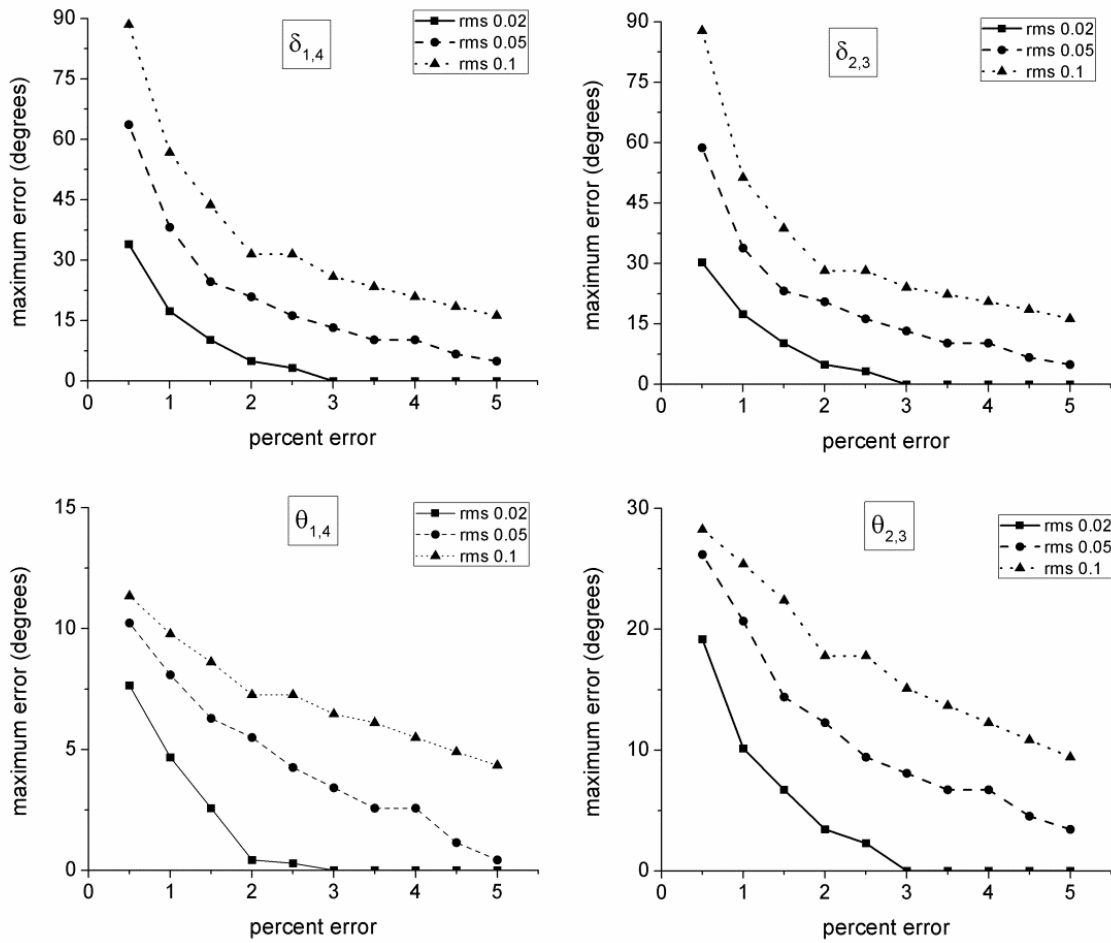


Figura 9: gráficas de los errores experimentales en los parámetros de los LCVRs para diferentes niveles de ruido en la medición de las intensidades y diferentes valores del error total rms requerido.

Conclusiones

En este trabajo se presentaron resultados de simulaciones de polarímetros de Mueller, incluyendo errores experimentales en el arreglo de los polarímetros y ruido de medición. También se incluyó el método de calibración de eigenvalores, también conocido como el método de Compain, y que se utiliza para corregir los errores en las mediciones. Encontramos los límites de errores experimentales y ruido de medición para que funciona correctamente el proceso de extracción de datos y el proceso de calibración, y las tolerancias en estos valores para obtener valores requeridos del error total rms en las matrices de Mueller recuperados. Sorprendentemente, las tolerancias sobre los ángulos de los ejes rápidos de los LCVRs son más estrictos que las tolerancias sobre los valores de las retardancias. Estos resultados pueden apoyar en el diseño de polarímetros.

Agradecimientos

Este trabajo fue realizado en colaboración con los siguientes personas: Dr. Oscar Gabriel Rodríguez Herrera, Dra. Martha Rosete Aguilar, Dr. Juan Manuel López Téllez, Dr. Omar Rodríguez Núñez, Oscar González Siu, Iván Montes González, Eusebio Aguilar Fernández. También, el trabajo fue apoyado por la DGAPA UNAM a través de proyectos PAPIIT TA100219, TA101020 and IG100121, y CONACYT a través de becas para estudiantes de maestría y doctorado.

Referencias

- [1] D. Goldstein, Polarized Light Second Edition, CRC Press, Boca Raton, (2003)
- [2] R.C. Jones, A new calculus for the treatment of optical systems, I. Description and Discussion of the Calculus, Journal of the Optical Society of America 31, 488-493 (1941).
- [3] R.C. Jones, A new calculus for the treatment of optical systems, IV Journal of the Optical Society of America, 32, 486-493 (1942).
- [4] G.G. Stokes, On the composition and resolution of streams of polarized light from different sources, Trans. Camb. Phil. Soc., 9, (1852), p. 399. Reimpreso en Mathematical and Physical Papers Volume 3, Cambridge University Press, London, 233 (1901).
- [5] F. Snik, J. Craven-Jones, M. Escuti, S. Fineschi, D. Harrington, A. D. Martino, D. Mawet, J. Riedi, and J. S. Tyo, An overview of polarimetric sensing techniques and technology with applications to different research fields,” Proceedings SPIE 9099, (2014), 90990B
- [6] X. Tan, X. Tu, K. Della Croce, G. Yao, H. Cai, N. Brock, S.Pau and R. Liang, Multi-wavelength quantitative polarization and phase microscopy, Biomedical Optics Express, 10(4), (2019), 1638-1648
- [7] B.H. Malik, G.L. Coté, Characterizing dual wavelength polarimetry through the eye for monitoring glucose, Biomedical Optics Express, 1(5), (2010), 1247-1258
- [8] D. Ivanov, T. Genova-Hristova, E. Borisova, L. Nedelchev, and D. Nazarova, Multiwavelength polarimetry of gastrointestinal ex vivo tissues for tumor diagnostic improvement, Proceedings SPIE 11047, (2019), 1104707
- [9] M. Kupinski, M. Boffety, F. Goudail, R. Ossikovski, A. Pierangelo, J. Rehbinder, J. Vizet and T. Novikova, Polarimetric measurement utility for pre-cancer detection from uterine cervix specimens, Biomedical Optics Express, 9(11), (2018), 5691-5702

- [10] J.S. Tyo, D.L. Goldstein, D.B. Chenault, J.A. Shaw, Review of passive imaging polarimetry for remote sensing applications, *Applied Optics*, 45(22), (2006), 5453-5469
- [11] W. Chen, S. Bai, D. Wang, H. Zhao, H. Sun, L. Yi, H. Zhao, D. Xie, J. Peltoniemi, Z. Li, Aerosol-induced changes in sky polarization pattern: potential hint on applications in polarimetric remote sensing, *International Journal of Remote Sensing*, 41(13), (2020), 4963-4980
- [12] B. P. Cumming, G. E. Schröder-Turk, S. Debbarma, and M. Gu, "Braggmirror-like circular dichroism in bio-inspired quadruple-gyroid 4srs nanostructures," *Light. Sci. & Appl.* **6**, e16192–e16192 (2017).
- [13] A. Y. Zhu, W. T. Chen, A. Zaidi, Y.-W. Huang, M. Khorasaninejad, V. Sanjeev, C.-W. Qiu, and F. Capasso, "Giant intrinsic chiro-optical activity in planar dielectric nanostructures," *Light. Sci. & Appl.* **7**, 17158– 17158 (2018).
- [14] A. Alvarez-Herrero, N. Uribe-Patarroyo, P. García Parejo, J. Vargas, R. L. Heredero, R. Restrepo, V. Martínez-Pillet, J. C. del Toro Iniesta, A. López, S. Fineschi, G. Capobianco, M. Georges, M. López, G. Boer, and I. Manolis, "Imaging polarimeters based on liquid crystal variable retarders: an emergent technology for space instrumentation," in *Proc. SPIE*(2011), Vol. 8160, p. 81600Y.
- [15] L. J. November, "Liquid crystal polarimeter: a solid state imager for solar vector magnetic fields," *Opt. Eng.* **34**, 1659 (1995).
- [16] V. Martínez Pillet, M. Collados, J. Sánchez Almeida, V. González, A. Cruz-Lopez, A. Manescau, E. Joven, E. Paez, J. Diaz, O. Feeney, V. Sánchez, G. Scharmer, and D. Soltau, "LPSP & TIP: Full Stokes Polarimeters for the Canary Islands Observatories," *ASP Conference Series*, 183, (1999), p. 264
- [17] G. Arfken, *Mathematical methods for physicists*, Academic Press Inc., New York, 1985, Capítulo 4
- [18] J.S. Tyo, Design of optimal polarimeters: maximization of signal to noise ratio and minimization of systematic errors, *Applied Optics*, 41(4), (2002), 619-630
- [19] A. de Martino, Y-K- Kim, E. García-Caurel, B. Laude and B. Drévilon, Optimized Mueller polarimeter with liquid crystals, *Optics Letters*, 28(8), (2003), 616-618
- [20] E. Compain, S. Poirier and B. Drevillon, "General and self-consistent method for the calibration of polarization modulators, polarimeters and Mueller-matrix ellipsometers", *App. Opt.*, **38**(16), 3490-3502 (1999)
- [21] N.C. Bruce, O.G. Rodríguez-Herrera, J.M. López-Téllez, O. Rodríguez-Nuñez, "Experimental limits for eigenvalue calibration in liquid-crystal Mueller-matrix polarimeters", *Optics Letters*, **43**, 2712-2715 (2018)

[22] N.C. Bruce, J.M. López-Téllez, O. Rodríguez-Nuñez, O.G. Rodríguez-Herrera, "Permitted experimental errors for optimized variable-retarder Mueller-matrix polarimeters", *Optics Express*, **26**, 13693-13704 (2018)

Materia cuántica en cavidades de alta reflectancia (Many-body CQED)

Santiago F. Caballero-Benítez

*Instituto de Física, LSCSC-LANMAC, Universidad Nacional Autónoma de México,
Apdo. Postal 20-364, México D. F. 01000, México.*

(Dated: November 15, 2021)

En este artículo se discuten algunos detalles del curso sobre “Materia cuántica en redes ópticas y cavidades de alta reflectancia (Many-body CQED)” de la escuela de verano de Física 2021 de la UNAM. Se describen conceptos útiles para la simulación cuántica análoga, así como se introducen detalles de modelos de materia cuántica en redes ópticas y en cavidades de alta reflectancia.

Las notas del curso son una introducción sobre la motivación de la simulación cuántica análoga y su relación con modelos típicos de estudio con la siguiente estructura:

- Redes ópticas y materia condensada
- El modelo de Bose-Hubbard
- Modelos en cavidades de alta reflectancia
- Conclusiones

I. REDES ÓPTICAS Y MATERIA CONDENSADA

A partir de los avances en el control de sistemas atómicos de manera experimental desde finales de la década de los 90, se ha logrado poder enfriar a temperaturas muy cercanas al cero absoluto (del orden de $10^{-6} - 10^{-9}\text{K}$) átomos en sistemas al alto vacío. En estas condiciones es posible atrapar a los átomos de muy baja energía en estructuras periódicas generadas por luz de láseres[1]. A estas estructuras periódicas se les denomina redes ópticas (“optical lattices”) [2]. Los átomos en estas redes ópticas se comportan de manera muy similar a cómo se comportan los electrones en un sólido. Esto da origen a la idea de utilizar sistemas atómicos como simuladores potenciales de materiales, donde a diferencia de un material real, todos los parámetros que dan forma a los modelos efectivos pueden ser controlados con alta precisión de manera externa experimentalmente[3, 4]. Desde el punto de vista teórico, esto abre la posibilidad a la exploración de nuevos modelos, que son experimentalmente realizables donde se pueden explorar mecanismos en búsqueda de nueva fenomenología y en varias instancias encontrar nuevas maneras en la materia se organiza en el régimen cuántico. Esto da lugar a nuevos estados de la materia, los cuales tienen el potencial de ser utilizados para el desarrollo de tecnología, ya sea en su realización en el sistema atómico o en un sistema análogo del estado sólido. Estas y otras aplicaciones de la materia cuántica forman parte de las “Tecnologías Cuánticas”. En ellas se busca utilizar efectos y propiedades de sistemas en el régimen cuántico para el desarrollo tecnológico[5].

Concretamente, la analogía entre los átomos en redes ópticas y el sistema electrónico se puede entender partir de la representación de los modelos de enlace fuerte electrones. Un modelo simple típico efectivo de electrones en un sólido es el modelo de Hubbard[6]:

$$\mathcal{H}^e = -t \sum_{\sigma \in \{\uparrow, \downarrow\}} \sum_{\langle i, j \rangle} (\hat{c}_{i, \sigma}^\dagger \hat{c}_{j, \sigma} + \text{H.c.}) + U \sum_i \hat{n}_{i, \uparrow} \hat{n}_{i, \downarrow} \quad (1)$$

donde $\hat{c}_{i, \sigma}^\dagger$ ($\hat{c}_{i, \sigma}$) es el operador de creación (aniquilación) de un electrón en un sitio i de la banda de conducción (c) con proyección de espín $\sigma \in \{\uparrow, \downarrow\}$. El símbolo $\langle i, j \rangle$ se refiere a los sitios primeros vecinos i, j . Los operadores de creación y aniquilación obedecen el algebra canónica de anti-conmutación tal que, $[\hat{c}_{i, \sigma}, \hat{c}_{j, \nu}^\dagger]_+ = \delta_{\sigma, \nu} \delta_{i, j}$, $[\hat{c}_{i, \sigma}^\dagger, \hat{c}_{j, \nu}^\dagger]_+ = 0$ y $[\hat{c}_{i, \sigma}, \hat{c}_{j, \nu}]_+ = 0$, donde $[A, B]_\pm = AB \pm BA$ es el anti-conmutador (conmutador). El operador del número de electrones en cada sitio con proyección σ de espín esta definido como: $\hat{n}_{\sigma, i} = \hat{c}_{i, \sigma}^\dagger \hat{c}_{i, \sigma}$. El segundo termino en (1) es conocido como el término de interacción de Hubbard[7]. El parámetro t es la amplitud de tunelamiento y U es la fuerza de interacción. El primer término de (1) tiene la interpretación de ser la energía cinética efectiva del sistema. En esta interpretación, los electrones se mueven de sitio en sitio en la red que forma el sólido por el potencial efectivo que se genera a partir de los potenciales moleculares efectivos, ver figura 2 (a). En una red óptica se tiene un potencial periódico con comportamiento similar dado por las ondas estacionarias generadas por los láseres, ver figura 2 (b). Al añadir la interacción electrónica (término con U) entender la dinámica de los electrones es uno de los problemas de mayor interés in la materia condensada, ya que se piensa que este modelo contiene los ingredientes esenciales para representar el fenómeno de superconductividad de alta temperatura crítica (high-Tc). Al día de hoy no existe una teoría definitiva que explique el mecanismo detrás de el fenómeno superconductividad en estos sistemas, llamados cupratos, y otros superconductores exóticos. El entender los elementos que controlan la high-Tc es de gran interés para el desarrollo de aplicaciones tecnológicas y su optimización[5].

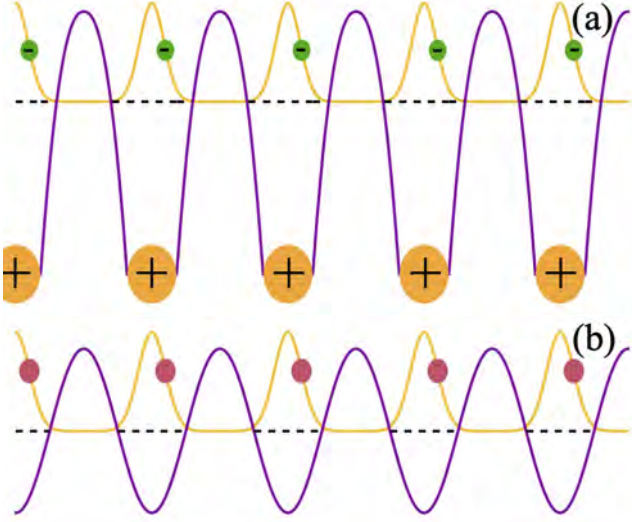


FIG. 1: (a) Esquema del potencial efectivo del modelo de enlace fuerte de electrones. (b) Esquema átomos ultrafríos en el potencial efectivo de una red óptica.

A. Redes ópticas clásicas

Las redes ópticas en los sistemas atómicos se puede realizar dado el efecto del campo eléctrico en los átomos, el efecto Stark [8]. Los átomos en un campo eléctrico tienen una respuesta con un momento dipolar $d_\nu^{(\pm)}$ dado por,

$$d_\nu^{(\pm)} = \sum_{\mu \in \{x,y,z\}} \alpha_{\nu\mu}(\omega_L) E_\mu^{(\pm)}(\mathbf{r}, t) \quad (2)$$

donde $\nu \in \{x, y, z\}$, ω_L es la frecuencia del laser, $\alpha_{\nu,\mu}(\omega_L)$ son los elementos de matriz del tensor de polarizabilidad complejo, \mathbf{r} es la posición y t el tiempo. De aquí se origina que el corrimiento del efecto Stark de corriente alterna (A.C.) cuadrático sirve para desintonizar con respecto a la frecuencia de resonancia atómica del sistema. El cambio en la energía que percibe el átomo es,

$$\Delta E(\mathbf{r}, t) = -2 \sum_{\nu, \mu \in \{x,y,z\}} \text{Re}[\alpha_{\nu\mu}(\omega_L) E_\nu^{(-)}(\mathbf{r}, t) E_\mu^{(+)}(\mathbf{r}, t)] \quad (3)$$

de aquí se deduce que el potencial efectivo, la red óptica, que ven los átomos es:

$$V_{\text{OL}} \approx \Delta E \propto \frac{I(\mathbf{r})}{\Delta_a} \quad (4)$$

donde $I(\mathbf{r}) \propto |E(\mathbf{r})|^2$ es la intensidad del laser y Δ_a es la desintonía con respecto a la frecuencia de resonancia atómica. Esto de manera efectiva es el cambio en la energía vista por los átomos debido a los láseres asumiendo una respuesta isotrópica del medio (los átomos).

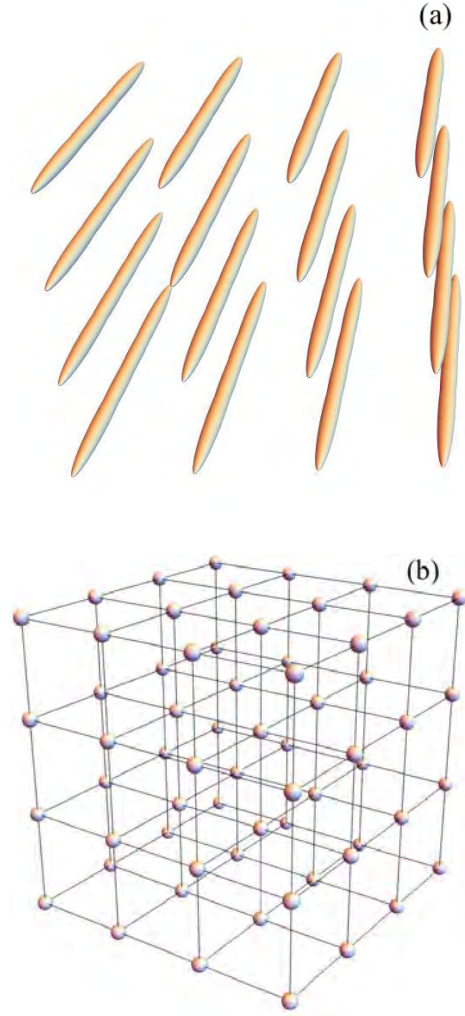


FIG. 2: (a) Red óptica de tubos o "cigarros", los haces contra-propagantes actúan sobre los ejes "x" y "z" en 3D. (b) Red óptica cúbica, los haces contra-propagantes actúan sobre los ejes "x", "y" y "z" en 3D.

Típicamente uno puede arreglar geoméricamente los láseres en configuraciones de pares de haces contra-propagantes en diferentes planos espaciales en una, dos, tres dimensiones e incluso otros planos efectivos adicionales arbitrarios[9–11]. Ejemplos para se muestran en la figura 2 (a) para arreglos de tubos o "cigarros" y (b) una red cúbica. Es posible generar todo tipo de redes periódicas, de Bravais e incluso cuasi-cristalinas con suficientes láseres. Para una red hiper-cúbica de dimensión d uno tiene para el potencial de la red óptica,

$$V_{\text{OL}} = \sum_{\nu=1}^d V_0 \sin^2(k_\nu x_\nu) \quad (5)$$

donde k_ν es el vector de onda de la luz en la dirección \hat{e}_ν , x_ν es la coordenada (i. e. x, y, z, \dots) tal que para para cada dimensión ν tenemos vectores unitarios \hat{e}_ν . Debido

a que el potencial efectivo que obtenemos se obtiene a sin tomar en cuenta aspectos cuánticos de la luz en el sistema y los campos de luz están siendo utilizados de manera efectiva paramétricamente, nos referiremos a este tipo de redes ópticas como “redes ópticas clásicas”.

II. EL MODELO DE BOSE-HUBBARD

El modelo de Hubbard de electrones (fermiones), tiene su contraparte en el sistema atómico en una red óptica clásica donde se pueden realizar experimentalmente sistemas fermiónicos (${}^6\text{Li}$, ${}^{40}\text{K}$) o sistemas bosónicos (${}^{87}\text{Rb}$). La primera realización experimental se realizó con átomos de Rubidio con estadística bosónica. En este caso el modelo de Bose-Hubbard[1, 2, 12] describe muy bien los experimentos[13]. El modelo se escribe como:

$$\mathcal{H}^b = -t \sum_{\langle i,j \rangle} (\hat{b}_i^\dagger \hat{b}_j + \text{H.c.}) + \frac{U}{2} \sum_i \hat{n}_i(\hat{n}_i - 1) - \mu \sum_i \hat{n}_i \quad (6)$$

donde \hat{b}_i^\dagger (\hat{b}_i) es el operador de creación (aniquilación) de un átomo en un sitio i con estadística bosónica. Los operadores de creación y aniquilación obedecen el algebra canónica de conmutación tal que $[\hat{b}_i, \hat{b}_j^\dagger]_- = \delta_{i,j}$, $[\hat{b}_i^\dagger, \hat{b}_j^\dagger]_- = 0$ y $[\hat{b}_i, \hat{b}_j]_- = 0$. El operador del número de átomos en cada sitio esta definido como: $\hat{n}_i = \hat{b}_i^\dagger \hat{b}_i$. Los parámetros del modelo son la amplitud de tunelamiento t , la fuerza de la interacción U y el potencial químico μ .

La modelo tiene dos límites simples de analizar. Para $t = 0$ (límite atómico) el Hamiltoniano es diagonal. Para llenado commensurado el estado base es el de mN átomos distribuidos cada uno en N sitios, donde m es un entero. El estado base es simplemente:

$$|\psi\rangle_{\text{MI-m}} = |m, m, \dots, m\rangle$$

En este caso, las fluctuaciones en el número de partículas, $\Delta n^2 = \langle \hat{n}_i^2 \rangle - \langle \hat{n}_i \rangle^2 = 0$. Los átomos están completamente localizados. Por ejemplo para una densidad por sitio $\rho = \langle \hat{n}_i \rangle = m = 1$, tenemos un átomo en cada sitio localizado, a este estado se le llama el estado de “aislante de Mott” m (Mott-Insulator, MI).

Por otro lado, para $U = 0$, el estado base es completamente des-localizado y presenta coherencia, en este caso se le llama estado “Superfluido” (SF). Las fluctuaciones en el número de partículas por sitio son diferentes de cero y en general pueden ser grandes, del orden del numero de partículas por sitio, $\Delta n^2 \sim \rho$. El estado se puede escribir como,

$$|\psi\rangle_{\text{SF}} \propto \sum_q \sqrt{\frac{M!}{q_1! q_2! \dots q_N!}} |q_1, q_2, \dots, q_N\rangle$$

donde q son las posibles combinaciones del conjunto de ocupaciones q_n en N sitios para M átomos. El cambio de entre estos dos estados al modificar los parámetros del modelo o en el sistema experimental es un escenario típico de una “transición de fase cuántica” (QPT) [14].

A. Aproximación de desacoplamiento, Teoría de perturbaciones y diagrama de fase

En esta sub-sección seguimos el tratamiento esbozado en [1, 15] para determinar el diagrama de fase con la aproximación de desacoplamiento y campo promedio.

Para poder tener una estimación de las fronteras de la transición de fase cuántica entre el estado superfluido y estado del aislante de Mott (Mott-Insulator)[12], es necesario realizar aproximaciones. Una aproximación que da buenos resultados cualitativos para este sistema es la aproximación de desacoplamiento en combinación con la aproximación de campo promedio. La aproximación consiste en desacoplar el producto de los operadores de creación y aniquilación en diferentes sitios, suponiendo que las correlaciones cuánticas entre ellos son despreciables. Adicionalmente se define un parámetro de orden, en este caso del orden superfluido. Este parámetro de orden se construye a partir del valor promedio de los valores de expectación en el estado base de los operadores de creación y aniquilación en el estado base. Esto es,

$$\hat{b}_i^\dagger \hat{b}_j \approx \langle \hat{b}_i^\dagger \rangle \hat{b}_j + \hat{b}_i^\dagger \langle \hat{b}_j \rangle - \langle \hat{b}_i^\dagger \rangle \langle \hat{b}_j \rangle \quad (7)$$

$$= \psi_i^* \hat{b}_j + \hat{b}_i^\dagger \psi_j - \psi_i^* \psi_j \quad (8)$$

donde $\psi_i^* = \langle \hat{b}_i^\dagger \rangle$ y $\psi_i = \langle \hat{b}_i \rangle$ para un sitio i . ψ es el parámetro de orden superfluido y su significado se aclarará a continuación.

Después de hacer la aproximación de desacoplamiento en el modelo de Bose Hubbard se puede obtener un Hamiltoniano efectivo por sitio suponiendo que el sistema es homogéneo y presenta invariancia translacional.

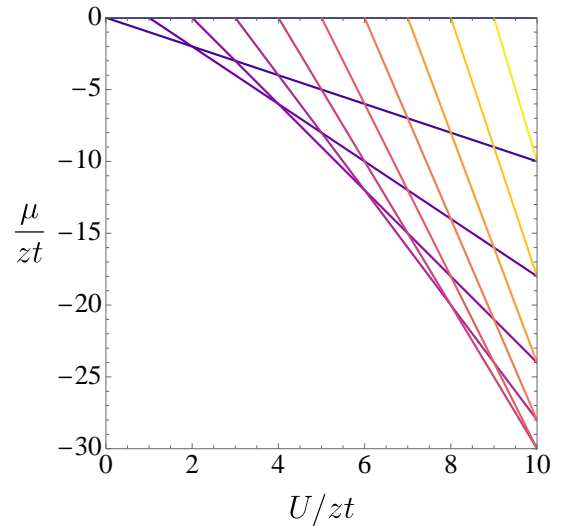


FIG. 3: Espectro de energías de los estados del Hamiltoniano no perturbado. Las curvas de los estados base como función de $\tilde{\mu}$ se intersectan en los puntos $\tilde{\mu}/\tilde{U} \in \mathcal{Z} \geq 0$

El Hamiltoniano por sitio efectivo es:

$$\tilde{\mathcal{H}} = -\psi(\hat{b}^\dagger + \hat{b}) + \psi^2 - \tilde{\mu}\hat{n} + \tilde{U}\hat{n}(\hat{n} - 1), \quad (9)$$

donde $\hat{n} = \hat{b}^\dagger\hat{b}$, $\psi = \langle \hat{b} \rangle$, $\tilde{\mu} = \mu/(zt)$, $\tilde{U} = U/(zt)$. En lo anterior hemos suprimido el índice por sitio en todas las cantidades y operadores. Además, hemos supuesto sin perder la generalidad para este caso que ψ es un número real y hemos supuesto que en promedio es el mismo a lo largo del todo el sistema. Es conveniente re-escribir el Hamiltoniano de la siguiente manera,

$$\tilde{\mathcal{H}} = \tilde{U}\tilde{\mathcal{H}}_0 + \psi\hat{V} + \psi^2, \quad (10)$$

donde $\hat{V} = (\hat{b} + \hat{b}^\dagger)$. Es conveniente re-escalar con respecto a \tilde{U} , entonces se obtiene:

$$\tilde{\mathcal{H}}_0 = -\frac{\tilde{\mu}}{\tilde{U}}\hat{n} + \hat{n}(\hat{n} - 1). \quad (11)$$

Adicionalmente, tenemos que $\hat{n}|n\rangle = n|n\rangle$. Debido a que ψ^2 es solo una constante, por el momento no es importante, siendo un corrimiento al espectro de energías. Los eigenvalores (valores propios) de $\tilde{\mathcal{H}}_0$ son:

$$\tilde{\mathcal{H}}_0|g\rangle = E_g|g\rangle \quad (12)$$

$$E_g = 0, \quad \frac{\tilde{\mu}}{\tilde{U}} < 0 \quad (13)$$

$$E_g = -\frac{\tilde{\mu}}{\tilde{U}}g + g(g - 1), \quad (g - 1) < \frac{\tilde{\mu}}{\tilde{U}} < g \quad (14)$$

Con la ayuda de la figura 3 del espectro se puede ver que la relación anterior es válida. Es posible definir una función simple $\tilde{g} = [\tilde{\mu}/\tilde{U} + 1]$, que da como resultado la parte entera de $[\cdot]$. Entonces, se tiene que

$$E_g = -\tilde{g}\tilde{\mu}/\tilde{U} + \tilde{g}(\tilde{g} - 1). \quad (15)$$

Utilizando teoría de perturbaciones independiente del tiempo no degenerada[?] y el resultado anterior, se obtiene,

$$E_g(\psi) = \tilde{U}E_g + \psi^2 + \psi \underbrace{\langle g|\hat{V}|g\rangle}_0 + \frac{2}{\tilde{U}} \sum_{n \neq g} \frac{|\langle g|\hat{V}|n\rangle|^2}{E_g - E_n} + O(\psi^3). \quad (16)$$

Se observa que la corrección de primer orden es cero y la corrección de segundo orden puede ser fácilmente calculada con ayuda de las siguientes definiciones,

$$\hat{b}|n\rangle = \sqrt{n}|n - 1\rangle \quad (17)$$

$$\hat{b}^\dagger|n\rangle = \sqrt{n + 1}|n + 1\rangle \quad (18)$$

En consecuencia, se obtiene lo siguiente,

$$\sum_{n \neq g} \frac{|\langle g|\hat{V}|n\rangle|^2}{E_g - E_n} = \left(\frac{\tilde{g} + 1}{E_g - E_{g+1}} + \frac{\tilde{g}}{E_g - E_{g-1}} \right) \quad (19)$$

Utilizando lo anterior se puede escribir lo siguiente,

$$E_g(\psi) = \tilde{U}E_g + \frac{a(\tilde{\mu}, \tilde{U})}{\tilde{U}}\psi^2 + O(\psi^4). \quad (20)$$

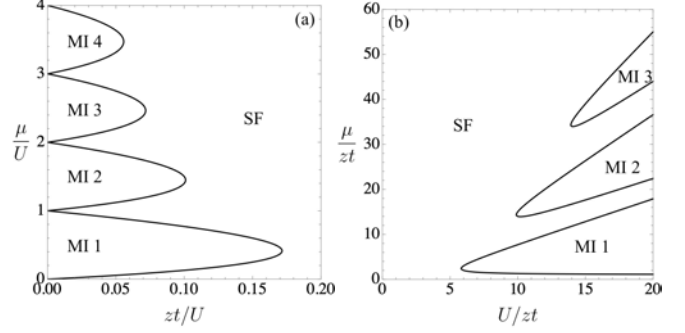


FIG. 4: Diagrama de fase del modelo de Bose-Hubbard en aproximación de campo promedio. (a) Presentación típica en materia condensada. (b) Presentación típica en Física atómica y en experimentos.

Donde la expresión anterior tiene la forma típica de un funcional de Landau. Uno puede identificar ψ con un parámetro de orden y concluir lo siguiente. Si $\psi \neq 0$ entonces a debe ser negativo para minimizar la energía. Si de manera contraria, a es positivo, entonces $\psi = 0$. Por lo tanto, los valores para los cuales esta transición sucede (la línea de transición) es controlada por el parámetro de orden y puede ser encontrada como una función de $\tilde{\mu}$ y de \tilde{U} . Se tiene que para $\psi \neq 0$ el sistema está en el estado superfluido y para $\psi = 0$ el sistema está en estado aislante de Mott. Para que suceda la transición se tiene que,

$$a(\tilde{\mu}, \tilde{U}) = \tilde{U} + \frac{\tilde{g} + 1}{\frac{\tilde{\mu}}{\tilde{U}} - \tilde{g}} + \frac{\tilde{g}}{-\frac{\tilde{\mu}}{\tilde{U}} + \tilde{g} - 1} = 0, \quad (21)$$

Se puede re-escribir la relación de arriba de manera conveniente como:

$$y = \frac{-\tilde{g}^2 + 2x\tilde{g} + \tilde{g} - x^2 - x}{1 + x}, \quad (22)$$

con $y = 1/\tilde{U}$ y $x = \tilde{\mu}/\tilde{U}$. Haciendo una gráfica de la función anterior se obtiene el diagrama de fase típico del modelo de Bose-Hubbard, ver la figura 4.

III. MODELOS EN CAVIDADES DE ALTA REFLECTANCIA

A. Esquemas de cavidades

Es posible experimentalmente colocar una cavidad de alta reflectancia en un experimento de átomos ultrafríos en redes ópticas clásicas[16–18]. Algunos típicos arreglos posibles con cavidades de alta reflectancia son los de el esquema de onda estacionaria (a) y onda viajera (b)[19], estos son mostrados en la figura 5. Como se puede apreciar en la figura es posible colocar la cavidad en diferentes ángulos con respecto a eje de la red óptica

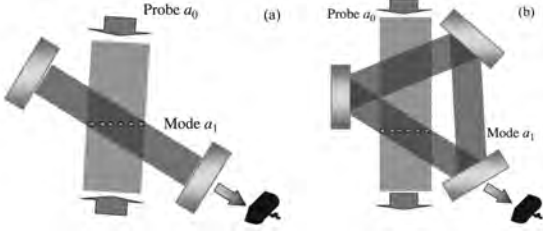


FIG. 5: Esquemas experimentales para sistemas de átomos ultrafríos en cavidades de alta reflectancia (Figuras de [19]). (a) Arreglo de cavidad de ondas estacionarias. (b) Arreglo de cavidad de ondas viajeras.

clásica, es posible colocar diversos detectores de los fotones que escapan la cavidad y se puede hacer incidir luz de bombeo en múltiples direcciones en principio. Adicionalmente, es posible tener más de una cavidad[20, 21] o tener una cavidad multi-modal[22].

B. Hamiltoniano de luz-materia cuántico

La materia ultrafría dentro de una cavidad de alta reflectancia con una red óptica clásica se puede representar con el siguiente Hamiltoniano de luz-materia en el régimen cuántico[23–25],

$$\mathcal{H} = \mathcal{H}^l + \mathcal{H}^{lm} + \mathcal{H}^m \quad (23)$$

donde \mathcal{H}^l es el Hamiltoniano de la luz, \mathcal{H}^m es la parte del Hamiltoniano que solo depende de la materia y \mathcal{H}^{lm} es la interacción entre la luz y la materia. A diferencia de las redes ópticas clásicas ahora la naturaleza cuántica de la luz es tomada en cuenta y esta afectará de manera significativa los procesos de correlación que habrá en la materia. Explícitamente los términos son:

$$\mathcal{H}^l = \sum_c \hbar\omega_c \hat{a}_c^\dagger \hat{a}_c, \quad (24)$$

$$\mathcal{H}^{lm} = \sum_{c,p,\sigma} \left(g_{pc\sigma}^* \hat{a}_c \hat{F}_{pc\sigma}^\dagger + g_{pc\sigma} \hat{a}_c^\dagger \hat{F}_{pc\sigma} \right) \quad (25)$$

y \mathcal{H}^m es el Hamiltoniano de Bose-Hubbard, Hubbard o incluso puede ser el de una mezcla Bose-Fermi. Los operadores de la luz son \hat{a}_c donde c es el sub-índice que denota los modos en la cavidad. Por otro lado el sub-índice p denota los modos de la luz bombeada al sistema y σ toma en cuenta la proyección de espín correspondiente. Los operadores de luz obedecen relaciones de conmutación bosónica. Los acoplamientos de la luz y la materia (i.e. proporcionales a las frecuencias de Rabi) están dados por $g_{pc\sigma}$. Los operadores \hat{F} denotan la proyección de los modos de luz con respecto a la materia dentro de la cavidad. Los modos que la luz imprimen a la materia pueden ser descompuestos en modos de densidad \hat{D} y modos de enlace \hat{B} de la siguiente

forma,

$$\hat{F}_{pc\sigma} = \hat{D}_{pc\sigma} + \hat{B}_{pc\sigma} \quad (26)$$

$$\hat{D}_{pc\sigma} = \sum_i J_{ii}^{pc\sigma} \hat{m}_{i\sigma}^\dagger \hat{m}_{i\sigma} \quad (27)$$

$$\hat{B}_{pc\sigma} = \sum_{\langle i,j \rangle} J_{ij}^{pc\sigma} (\hat{m}_{i\sigma}^\dagger \hat{m}_{j\sigma} + \text{H.c.}) \quad (28)$$

donde los operadores $\hat{m}_{i\sigma}^\dagger$ pueden corresponder a fermiones \hat{f}^\dagger o bosones \hat{b}^\dagger . Además, las constantes J_{ii} son los traslapes de modos de densidad (términos en sitio) y las constantes J_{ij} son los traslapes de modos de enlace (términos a primeros vecinos). Estas constantes de estructura están dadas por las siguientes integrales:

$$J_{ij}^{pc\sigma} = \int w(\mathbf{x} - \mathbf{x}_i) u_{c\sigma}^*(\mathbf{x}) u_{p\sigma}(\mathbf{x}) w(\mathbf{x} - \mathbf{x}_j) d^n x \quad (29)$$

donde $w(\mathbf{x})$ son las funciones de Wannier de la red óptica clásica en la que están definidos los operadores de materia que conforman \mathcal{H}^m . Tomando en cuenta que las constantes de estructura típicamente forman un conjunto finito de valores posibles en amplitud con patrones espaciales, es posible definir operadores de modos inducidos por luz tal que,

$$\hat{D}_\sigma = \sum_\varphi J_{\varphi,\sigma}^D \hat{N}_{\varphi,\sigma} \quad (30)$$

$$\hat{B}_\sigma = \sum_\varphi J_{\varphi,\sigma}^B \hat{S}_{\varphi,\sigma} \quad (31)$$

con φ los modos inducidos por luz (el conjunto de diferentes valores de J_{ij}) y los siguientes operadores para estos modos:

$$\hat{N}_{\varphi,\sigma} = \sum_{i \in \varphi} \hat{m}_{i,\sigma}^\dagger \hat{m}_{i,\sigma} \quad (32)$$

$$\hat{S}_{\varphi,\sigma} = \sum_{\langle i,j \rangle \in \varphi} (\hat{m}_{i,\sigma}^\dagger \hat{m}_{j,\sigma} + \text{H.c.}) \quad (33)$$

Típicamente las funciones de estructura de los modos inducidos por la luz tienen son de la forma (en 1D) $u_{p/c}(x) \propto \cos(2\pi qx/\lambda)$ para cavidades de onda estacionaria, donde q es una constante que depende del arreglo geométrico particular de los haces y espejos en el sistema. Explícitamente para diferentes configuraciones se pueden encontrar en [26]. El problema completo con grados de libertad de luz y materia es muy difícil de resolver. De ahí que una aproximación habitual es la de eliminar adiabáticamente la luz.

C. Modelos efectivos de materia cuántica en cavidades de alta reflectancia.

En el marco de rotación de la luz es útil definir la desintonía de la frecuencia de la cavidad con la luz de bombeo que se hace incidir en el sistema, $\Delta_c = \omega_c - \omega_p$,

donde $\omega_{c/p}$ son las frecuencias de la cavidad y el haz de bombeo. Por simplicidad en lo que sigue consideraremos una cavidad mono-modal y un haz de bombeo. Además se considera de manera fenomenológica una κ que es la tasa de pérdida de fotones de la cavidad. A partir de las ecuaciones de movimiento de los operadores de la luz y la materia, se considera el límite del estado estacionario de la luz. Esto es el límite donde

$$\left\langle \frac{d\hat{a}}{dt} \right\rangle \stackrel{!}{=} 0 \quad \rightarrow \text{límite adiabático} \quad (34)$$

a partir de esta condición, en el límite $\kappa \ll |\Delta_c|$, que significa que la pérdida de fotones en el sistema es pequeña o equivalentemente que la cavidad es muy buena, se puede aproximar cualitativamente,

$$\hat{a} \sim \frac{g\hat{F}}{\Delta_c + i\kappa} \quad (35)$$

Los detalles de diferentes maneras de realizar la aproximación o un tratamiento equivalente se encuentran en [27, 28]. El tratamiento a detalle de la eliminación adiabática lleva al siguiente Hamiltoniano efectivo sólo con grados de libertad de materia (sin espín),

$$\mathcal{H}_{\text{eff}}^m = \mathcal{H}^m + \frac{\hbar\Delta_c|g|^2}{2N_s(\Delta_c^2 + \kappa^2)}(\hat{F}^\dagger\hat{F} + \hat{F}\hat{F}^\dagger) \quad (36)$$

donde N_s es el número de sitios en la red óptica clásica que describe el sistema. Se debe notar que el número de átomos en el sistema es del orden del número de sitios. La luz dentro de la cavidad media una interacción en la materia de largo alcance que tiene estructura dictada por como se inyecta la luz al sistema y el arreglo geométrico de la cavidad, así como la red óptica.

Bosones sin espín. A partir del modelo efectivo en (36) es posible considerar el límite donde solamente acoplamiento en las densidades sucede. En la configuración donde el potencial de la red óptica clásica es lo suficientemente profundo $J_{ij} \approx 0$ para primeros vecinos y bombeando la luz a 90° con respecto al eje de la cavidad, $J_{ii} = J_D(-1)^i$. Entonces, se puede escribir el siguiente modelo efectivo,

$$\mathcal{H} = \mathcal{H}^b + \frac{g_{\text{eff}}}{N_s} \left[\sum_{\nu} (\hat{n}_{O,\nu} - \hat{n}_{E,\nu}) \right]^2, \quad (37)$$

donde ν corre sobre la mitad de los sitios y $g_{\text{eff}} \sim \hbar\Delta_c J_D^2 |g|^2 / (\Delta_c^2 + \kappa^2)$. Este modelo efectivo ha sido realizado experimentalmente [16, 17]. Se encuentra que además de existir las fases de la materia del estado aislante de Mott y estado superfluido, el sistema presenta un estado de aislante de Densidad (DW) con simetría Z_2 rota (Par-Impar) y uno con estado superfluido y de DW coexistiendo, el estado Supersólido (SS). El estado DW puede ser visualizado con el orden que sucede en un tablero de ajedrez donde los átomos se localizan en los sitios pares o en los sitios impares. Cálculos de este

modelo efectivo son consistentes con estos resultados [23–25, 29].

En el caso de que las amplitudes J_{ij} para $\langle i, j \rangle$ dominen el comportamiento por condiciones geométricas [26], estados con “orden de enlace” (bond order) son posibles. Esto da lugar a dímeros espaciales donde las fases relativas de las ondas de materia por sitio se auto-organizan [23–25].

Fermiones con espín. En este caso el Hamiltoniano de materia \mathcal{H}^m se puede escribir como,

$$\mathcal{H}^f = - \sum_{\sigma} \sum_{\langle i, j \rangle} t_{ij} \left(\hat{f}_{i\sigma}^\dagger \hat{f}_{j\sigma} + \text{H.c.} \right) + U_f \sum_i \hat{n}_{i\uparrow} \hat{n}_{i,\downarrow} \quad (38)$$

qué es el Hamiltoniano de Fermi-Hubbard, pero ahora en lugar de electrones tenemos átomos \hat{f} con estadística fermiónica y dos proyecciones de espín efectivas. En este caso, se puede considerar que la luz se acopla dependiendo de la polarización de la luz, derecha o izquierda R/L tal que: $a_{p,\theta} \propto (a_{p,L} e^{i\theta} + a_{p,R} e^{-i\theta})$. Esto genera el siguiente Hamiltoniano efectivo, eliminando adiabáticamente la luz [30],

$$\mathcal{H}_{\text{eff}}^f = \mathcal{H}^f + \frac{g_{\text{eff}}}{2N_s} (\hat{D}_\theta^\dagger \hat{D}_\theta + \hat{D}_\theta \hat{D}_\theta^\dagger) \quad (39)$$

En (39), se ha considerado que la red óptica es lo suficientemente profunda, se bombea luz a 90° y solo términos de densidad son relevantes,

$$\hat{D}_\theta = \sum_j (-1)^j (\hat{n}_{j,\uparrow} e^{i\theta} + \hat{n}_{j,\downarrow} e^{-i\theta}) \quad (40)$$

donde θ es el ángulo de polarización. Este Hamiltoniano puede soportar además de estados de pares superfluidos (SF), aislantes anti-ferromagnéticos (AFI) y estados de pares con simetría Z_2 rota formando una DW. El estado con DW además puede tener fracción superfluida de pares no nula, dando lugar a al estado de “onda de densidad de pares” (pair density wave PDW). Es posible hacer ingeniería de estados manipulando el ángulo de polarización, el factor de llenado y explorar diferentes estructuras en la competencia de estados SF, AFI, PDW y DW. Además es posible manipular el diagrama de fases generando estructuras similares a las de sistemas que presentan high-Tc, en la línea crítica de $T = 0$ [30].

Bosones con espín. Recientemente, el Hamiltoniano (37) ha sido extendido y estados con propiedades magnéticas no triviales han sido explorados [31].

IV. CONCLUSIONES

La luz en el límite cuántico induce interacciones efectivas que cambian significativamente el panorama de las fases de materia cuántica de muchos cuerpos en el sistema. Es posible diseñar por efectos de retro-acción por cavidades estados de materia cuántica emergente con correlaciones fuertes. Usando las propiedades de la luz

es posible inducir la competencia de ordenes en sistemas de materia cuántica y utilizarlos para hacer simulación cuántica avanzada. Los potenciales cuánticos inducidos

por la luz ofrecen una oportunidad única para innovar buscando nuevos sistemas de control cuántico e ir más allá de análogos de sistemas de materia condensada.

-
- [1] M. Lewenstein, A. Sampera, and V. Ahufinger. *Ultracold atoms in optical lattices: Simulating Quantum Many-Body Systems*. Oxford University Press (2012).
- [2] Jaksch, D., Bruder, C., Cirac, J. I., Gardiner, C. W. and Zoller, P. Cold Bosonic Atoms in Optical Lattices. *Phys. Rev. Lett.* **81**, 3108 (1998).
- [3] I. Bloch, J. Dalibard, and J. Sylvain. *Nat. Phys.* **8**, 267 (2012).
- [4] F. Schäfer, T. Fukuhara, S. Sugawa, Y. Takasu and Y. Takahashi. Tools for Quantum simulation with ultracold atoms in optical lattices. *Nat. Rev. Phys.* **2**, 411 (2020).
- [5] E. Altman, et. al. Quantum Simulators: Architectures and Opportunities. *Phys. Rev. X Quantum* **2**, 017003 (2021)
- [6] P. Coleman. *Introduction to Many-body Physics*. Cambridge University Press (2015).
- [7] J. Hubbard, Electron correlations in narrow energy bands. *Proc. R. Soc. A* **276**, 238 (1963).
- [8] C. Foot. *Atomic Physics*. Oxford University Press (2005).
- [9] I. Bloch, Ultracold quantum gases in optical lattices. *Nat. Phys.* **1**, 23 (2005).
- [10] O. Boada, A. Celi, J. I. Latorre, and M. Lewenstein, Quantum Simulation of an Extra Dimension. *Phys. Rev. Lett.* **108**, 133001 (2012).
- [11] D. I. Tsomokos, S. Ashhab, and F. Nori. Using superconducting qubit circuits to engineer exotic lattice systems. *Phys. Rev. A* **82**, 052311 (2010).
- [12] M. P. A Fisher., P. B. Weichman, G. Grinstein, and D. S. Fisher, Boson localization and the superfluid-insulator transition. *Phys. Rev. B* **40**, 546 (1989).
- [13] M. Greiner, O. Mandel, T. Esslinger, T. W. Hänsch and I. Bloch, Quantum phase transition from a superfluid to a Mott insulator in a gas of ultracold atoms. *Nature* **415**, 39 (2002).
- [14] S. Sachdev. *Quantum Phase Transitions*, Second Edition. Cambridge University Press (2011).
- [15] D. van Oosten, P. van der Straten, and H.T. C. Stoof, Quantum phases in an optical lattice. *Phys. Rev. A* **63**, 053601 (2001).
- [16] J. Klinder, H. Kekler, M. Reza Bakhtiari, M. Thorwart, and A. Hemmerich, Observation of a Superradiant Mott Insulator in the Dicke-Hubbard Model. *Phys. Rev. Lett.* **115**, 230403 (2015).
- [17] R. Landig, et. al. Quantum phases from competing short- and long-range interactions in an optical lattice. *Nature* **532**, 476 (2016).
- [18] L. Hrubya, N. Dogra, M. Landini, T. Donner, and T. Esslinger, Metastability and avalanche dynamics in strongly correlated gases with long-range interaction, *PNAS* **115** (13) 3279-3284 (2018).
- [19] I. B. Mekhov and H. Ritsch, Quantum Optics with Quantum Gases. *Laser Phys.* **19**, 610 (2009).
- [20] J. Léonard, A. Morales, P. Zupancic, T. Esslinger, and T. Donner. Supersolid formation in a quantum gas breaking a continuous translational symmetry. *Nature* **543**, 87 (2017).
- [21] A. Morales, P. Zupancic, J. Léonard, T. Esslinger, and T. Donner. Coupling two order parameters in a quantum gas. *Nat. Mat.* **17**, 686 (2018).
- [22] A. J. Kollár, A. T. Papageorge, K. Baumann, M. A. Armen, and B. L. Lev, An adjustable-length cavity and Bose-Einstein condensate apparatus for multimode cavity QED. *New J. Phys.* **17**, 043012 (2015).
- [23] S. F. Caballero-Benitez, and I. B. Mekhov, Quantum optical lattices for emergent many-body phases of ultracold atoms. *Phys. Rev. Lett.* **115**, 243604 (2015).
- [24] S. F. Caballero-Benitez, G. Mazzucchi, and I. B. Mekhov, Quantum simulators based on the global collective light-matter interaction. *Phys. Rev. A* **93**, 063632 (2016).
- [25] S. F. Caballero-Benitez, and I. B. Mekhov, Bond order via light-induced synthetic many-body interactions of ultracold atoms in optical lattices. *New J. Phys.* **18**, 113010 (2016).
- [26] W. Kozłowski, S. F. Caballero-Benitez, and I. B. Mekhov, Probing Matter-Field and Atom-Number Correlations in Optical Lattices by Global Nondestructive Addressing. *Phys. Rev. A* **92**, 013613 (2015).
- [27] C. Maschler, I. B. Mekhov, and H. Ritsch, Ultracold atoms in optical lattices generated by quantized light fields. *Eur. Phys. J. D* **46**, 545-560 (2008).
- [28] S. F. Caballero-Benitez, and I. B. Mekhov, Quantum properties of light scattered from structured many-body phases of ultracold atoms in quantum optical lattices. *New J. Phys.* **17** 123023 (2015).
- [29] N. Dogra, F. Brennecke, S. D. Huber and T. Donner, Phase transitions in a Bose-Hubbard model with cavity-mediated global-range interactions, *Phys. Rev. A* **94**, 023632 (2016).
- [30] A. Camacho-Guardian, R. Paredes, and S. F. Caballero-Benitez, Quantum Simulation of Competing Orders with Fermions in Quantum Optical Lattices. *Phys. Rev. A* **96**, 051602(R) (2017)
- [31] K. Lozano-Mendez, A. H. Casares and S. F. Caballero-Benitez. Spin Entanglement and Magnetic Competition via Long-range Interactions in Spinor Quantum Optical Lattices, arXiv:2011.07765 (2020)

Fuerzas y calor a la nanoescala

S. G. Castillo-López, R. Esquivel-Sirvent, G. Pirruccio, and C. Villarreal
*Instituto de Física, Universidad Nacional Autónoma de México,
Apartado Postal 20364, México 01000, México.*

INTRODUCCIÓN

Los estudios de la radiación térmica de los materiales han jugado un papel importante en la historia de la física. Basta mencionar que la mecánica cuántica se originó a partir de los intentos de explicar los resultados experimentales paradójicos relacionados con la radiación del cuerpo negro. Cotidianamente nos topamos con ejemplos tales como aquella proveniente como luz solar, o la emitida por los hornos de pan, los hornos de fundición, la lava incandescente, la radiación de fondo de 3 K, y aún por el cuerpo humano.

Hacia finales del siglo XIX, el espectro de radiación del cuerpo negro representaba una incógnita, ya que si se aplicaban los conceptos básicos de la física clásica, no se reproducía el espectro observado. Si se considera que la luz es emitida por las paredes de una cavidad formada por radiadores elementales representados por osciladores armónicos, entonces el espectro de energía, $\rho(\omega)$, para radiación de frecuencia angular ω , puede expresarse como el producto de la densidad de modos electromagnéticos del sistema, por la energía promedio E de cada modo. Si suponemos que el sistema se encuentra en equilibrio termodinámico entonces el teorema de equipartición nos dice que $E = k_B T$, de manera que:

$$\rho(\omega) = \frac{\omega^2}{2\pi^2 c^3} k_B T, \quad (1)$$

donde c es la velocidad de la luz, y k_B la constante de Boltzmann. Esta expresión, conocida como distribución de Rayleigh-Jeans, presenta diversos problemas, ya que conduce a una cantidad divergente en la integración $u = \int_0^\infty d\omega \rho(\omega)$ para obtener la densidad de energía total u . Este hecho fue conocido como la catástrofe del ultravioleta. Por otro lado, y fundamentalmente, no concuerda con el experimento a altas frecuencias.

En los albores del siglo XX, Max Planck, quien era experto en termodinámica, se abocó al estudio de estos problemas y encontró una solución revolucionaria que cambió profundamente las concepciones físicas aceptadas en esa época: el espectro observado experimentalmente se reproducía suponiendo que cada uno de los osciladores elementales sólo podía emitir energía en forma de paquetes cuantizados de magnitud $\hbar\omega$, donde la cantidad \hbar se denomina como la constante de Planck (dividida por 2π).

Esa hipótesis lo condujo a la distribución:

$$\rho(\omega) = \frac{\omega^2}{\pi^2 c^3} \frac{\hbar\omega}{e^{\hbar\omega/k_B T} - 1}, \quad (2)$$

conocida ahora como la distribución de Planck. Con ello, hacia la navidad de 1900 tuvo inicio la mecánica cuántica.

Una consecuencia inmediata de la distribución de Planck es que su integral sobre las frecuencias es finita, de modo que la densidad de energía del cuerpo negro está dada por

$$u = \frac{\pi^2 k_B^4 T^4}{15 \hbar^3 c^3}, \quad (3)$$

lo que permite determinar a su vez el flujo radiativo de calor por unidad de área de un cuerpo negro, $S = uc/4$ es decir, la ley de Stefan-Boltzman:

$$S = \frac{\pi^2 k_B^4 T^4}{60 \hbar^3 c^2}, \quad (4)$$

Esta relación será importante en lo que sigue, ya que se asocia a las ondas electromagnéticas propagantes que se emiten por un cuerpo aislado en equilibrio térmico a temperatura T ; sin embargo, en presencia de un cuerpo cercano, la interacción mutua mediada por sus respectivos campos electromagnéticos modifica esta expresión, ya sea porque el segundo cuerpo esté a una temperatura diferente, o porque se encuentre muy cercano. En este último caso, puede ocurrir lo que en óptica se conoce como reflexión total interna, asociada a ondas electromagnéticas evanescentes que decaen exponencialmente en la superficie de un medio, pero las cuales pueden transmitirse a otro que se encuentre a una distancia suficientemente pequeña L . Este hecho ha provocado el interés reciente en el estudio de la transferencia de calor radiativa a microescala y nanoescala, lo que ha planteado ha planteado nuevos desafíos, algunos de los cuales discutiremos en la sección siguiente.

Por otro lado, al poco tiempo de proponer su distribución, Planck se percató, al tratar de recuperar la distribución de Rayleigh-Jeans como el límite de alta temperatura de su distribución, que debía modificar su propuesta original introduciendo una contribución extra independiente de la temperatura, $\hbar\omega/2$, lo que condujo a una nueva distribución dada por

$$\rho(\omega) = \frac{\omega^2}{\pi^2 c^3} \left[\frac{\hbar\omega}{e^{\hbar\omega/k_B T} - 1} + \frac{\hbar\omega}{2} \right]. \quad (5)$$

Es fácil ver, mediante un desarrollo en serie de Taylor, que en el límite $\hbar\omega/k_B T \ll 1$ la presencia de dicha contribución conduce a la distribución clásica dada por la Ec.(1). Si bien el término $\hbar\omega/2$, se introdujo de manera formal, tiene consecuencias fundamentales. Ahora se sabe que está asociado a campos electromagnéticos derivados de fluctuaciones cuánticas de la materia que persisten aún a una temperatura $T = 0$ K, y son independientes de las fluctuaciones térmicas; por tanto, se le denomina como energía de punto cero o energía del vacío. Concluimos entonces que en la descripción física de los fenómenos radiativos, en general debemos tomar en cuenta ambas contribuciones. En la sección sobre fuerzas de Casimir examinaremos algunas de las manifestaciones físicas derivadas de la energía del vacío.

TRANSFERENCIA RADIATIVA DE CALOR DE CAMPO CERCANO

Resulta intuitivamente claro que el flujo radiativo de calor por unidad de área entre dos cuerpos negros separados por una distancia $L \gg \lambda_T = \hbar c/k_b T$ está determinado por la ley de Stefan-Boltzmann

$$S = \frac{\pi^2 k_B^4}{60 \hbar^3 c^2} (T_1^4 - T_2^4), \quad (6)$$

donde T_1 y T_2 son las temperaturas de los cuerpos 1 y 2, respectivamente. Para ello, cada uno de los cuerpos debe encontrarse en un estado de equilibrio térmico local. Como ya habíamos mencionado, en este caso límite, la transferencia de calor entre los cuerpos está determinada por ondas electromagnéticas propagantes irradiadas por los cuerpos y no depende de la separación L . Dichas ondas existen siempre en el exterior de cualquier cuerpo debido a fluctuaciones térmicas y cuánticas de la densidad de corriente dentro del mismo. Hasta hace relativamente poco tiempo, la parte de campo cercano evanescente de la radiación electromagnética generalmente se ignoraba porque no juega ningún papel en las propiedades de campo lejano de la emisión de ondas propagantes. Sin embargo, para dos sólidos separados por una distancia $L < T$, donde $\lambda_T = \hbar c/k_b T$, es la longitud de onda térmica, la transferencia de calor puede aumentar en muchos órdenes de magnitud debido a las ondas electromagnéticas evanescentes.

La teoría sobre la transferencia radiativa de calor de campo cercano originado a partir de campos electromagnéticos excitados térmicamente dentro de los materiales fue desarrollada por Rytov. Esta teoría tiene la ventaja de que para un sistema formado por placas paralelas, las consecuencias físicas de dichos campos fluctuantes se expresan en términos de coeficientes de reflexión para ondas electromagnéticas con polarización S (transversal eléctrica) y P (transversal magnética). En este sistema,

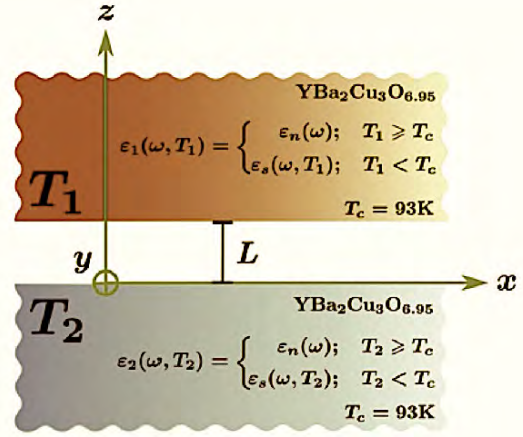


FIG. 1: Configuración determinada por dos placas paralelas semi-infinitas de YBCO separadas por una distancia L

la transferencia de calor radiativo entre las dos placas tiene contribuciones tanto de ondas propagantes como evanescentes. Si la utilizamos las ecuaciones de Maxwell para calcular el vector de Poynting, en donde se satisfacen las respectivas con condiciones de contorno entre las placas, el flujo de calor total resultante es

$$Q_T(L, T_1, T_2) = \int_0^\infty d\omega S_\omega(\omega, L, T_1, T_2), \quad (7)$$

donde S_ω describe el flujo de calor espectral de acuerdo con la expresión,

$$S_\omega(\omega, L, T_1, T_2) = [\Theta(\omega, T_1) - \Theta(\omega, T_2)] \sum_{j=p,s} \int \frac{d\beta d\beta}{(2\pi)^2} \left[\tau_j^{\text{prop}}(\omega, \kappa, L) + \tau_j^{\text{evan}}(\omega, \kappa, L) \right]. \quad (8)$$

Aquí, $\Theta(\omega, T) = \hbar\omega / [\exp(\hbar\omega/k_B T) - 1]$ es la energía promedio del oscilador de Planck, mientras que T_1 and T_2 son las temperaturas absolutas del cuerpo emisor y el receptor, con $T_1 > T_2$. Por otro lado, β y κ son las componentes del vector de onda paralelo y perpendicular a las superficies de las placas, respectivamente, y están relacionados mediante $\kappa = \sqrt{\omega^2/c^2 - \beta^2}$ en el vacío entre las placas y por $\kappa_i = \sqrt{\epsilon_i \omega^2/c^2 - \beta^2}$ dentro del medio caracterizado por la función dieléctrica ϵ_i , con $i = 1, 2$.

En la ecuación (8), la suma considera la contribución tanto de las ondas S y P mediante los coeficientes de transmisión τ_p y τ_s , los cuales tienen una expresión diferente para ondas propagantes ($\beta < \omega/c$) y evanescentes ($\beta > \omega/c$), como sigue

$$\tau_{j=p,s}^{\text{prop}}(\omega, \kappa, L) = \frac{(1 - |r_j^{(1)}|^2)(1 - |r_j^{(2)}|^2)}{|1 - r_j^{(1)} r_j^{(2)} \exp(2i\kappa L)|^2} \quad (9)$$

$$\tau_{j=p,s}^{\text{evan}}(\omega, \kappa, L) = \frac{4\text{Im}(r_j^{(1)})\text{Im}(r_j^{(2)}) \exp(-2|\kappa|L)}{|1 - r_j^{(1)} r_j^{(2)} \exp(-2|\kappa|L)|^2}, \quad (10)$$

TRANSFERENCIA RADIATIVA DE CALOR ENTRE SUPERCONDUCTORES

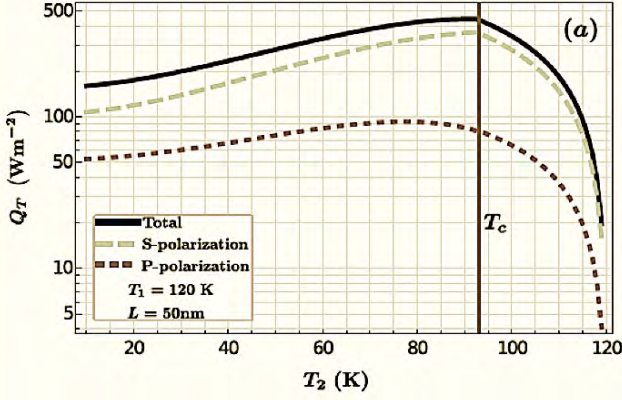


FIG. 2: Calor transmitido, Q , para una configuración de placas separadas por una distancia $L = 50\text{nm}$, suponiendo que una de ellas se mantiene una temperatura fija $T_1 = 120\text{ K}$, mientras que la temperatura de la segunda varía entre $T = 10\text{K}$ y $T = 120\text{ K}$.

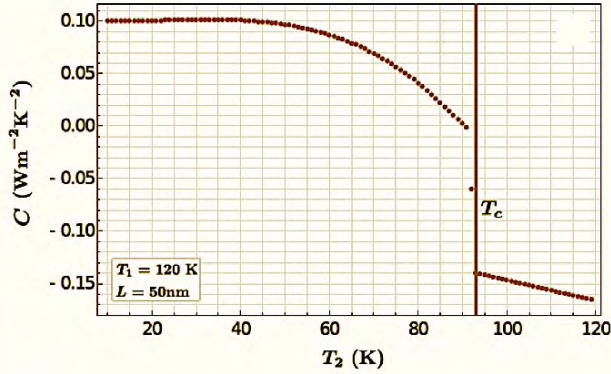


FIG. 3: Equivalente al calor específico del sistema, representado por la derivada dQ/dT , para una configuración de placas separadas por una distancia $L = 50\text{nm}$, suponiendo que una de ellas se mantiene una temperatura fija $T_1 = 120\text{ K}$, mientras que la temperatura de la segunda varía entre $T = 10\text{ K}$ y $T = 120\text{ K}$.

donde r_j^i son los coeficientes de reflexión de Fresnel entre la interfase i y el vacío. Para la polarización S este coeficiente tiene la forma

$$r_s^i = \frac{\kappa - \kappa_i}{\kappa + \kappa_i} \quad (11)$$

mientras que para la polarización P

$$r_p^i = \frac{\varepsilon_i \kappa - \kappa_i}{\varepsilon_i \kappa + \kappa_i} \quad (12)$$

Los autores de estas notas hemos explorado las consecuencias de la transferencia radiativa de calor entre superconductores a escalas nanométricas. La superconductividad es relevante en una variedad de campos que van desde la física de altas energías hasta la tecnología de microondas. La conducción sin pérdidas es atractiva porque permite un aumento de la eficiencia de dispositivos como consecuencia de la reducción del ruido y el menor consumo de energía. Los materiales superconductores están en el corazón de varias tecnologías, como los SQUIDS, metamateriales, o detectores de fotones, entre otros. También son adecuados para crear qubits en la información cuántica. En particular, para todos aquellos dispositivos que experimentan miniaturización, la proximidad de varios elementos superconductores operados por encima o por debajo de su temperatura crítica merece consideración. Los elementos superconductores muy compactos pueden presentar cambios en sus prestaciones debido al acoplamiento térmico. En este contexto, nosotros nos hemos enfocado en el estudio de cerámicas superconductoras de alta temperatura crítica, T_c . En estos materiales, T_c es del orden de 100 K y hemos considerado un caso prototípico, el $\text{YBa}_2\text{Cu}_3\text{O}_{7-\delta}$ (YBCO), con $T_c = 93\text{ K}$ en una configuración representada por dos placas paralelas separadas por una distancia L , de acuerdo a la figura 1.

La respuesta óptica del YBCO está determinada por su respuesta dieléctrica, $\varepsilon(\omega, T)$, la cual es diferente para temperaturas por encima y por debajo de la crítica. En el caso $T > T_c$, el material se comporta como un conductor normal, donde la corriente se transporta con pérdidas disipativas determinadas por la tasa de relajación electrónica γ_0 , mientras que en el caso $T < T_c$, el material se convierte en superconductor y esto está asociado al límite $\gamma_0 \rightarrow 0$. Entonces, en el caso normal la respuesta dieléctrica está dada por una expresión de Drude-Lorentz

$$\varepsilon_n(\omega) = \varepsilon_\infty - \frac{\omega_{pn}^2}{\omega^2 + i\gamma_p\omega} - \frac{\Omega_{mir}^2}{\omega^2 - \omega_{mir}^2 + i\Gamma_{mir}\omega}, \quad (13)$$

donde ε_∞ representa una contribución de altas frecuencias, el segundo término, tipo Drude, está asociado a las cargas conductoras y el tercero, de Lorentz, a electrones ligados en el mediano infrarrojo. Aquí, la frecuencia de plasma $\omega_p^2 = 4\pi n e^2/m$, donde n es la densidad de carga de partículas con carga e y masa m . Por otro lado, en el caso superconductor, de acuerdo al modelo de los dos fluidos, la densidad de carga se expresa como la suma $n = n_s(T) + n_n(T)$, donde $n_s(T)$ y $n_n(T)$ representan las fracciones de cargas en el estado superconductor y normal, respectivamente, la cuales varían con la temperatura. A temperatura cero, todas las cargas son superconductoras, es decir, $n_s(0) = n$. Como con-

secuencia, el modelo de dos fluidos conduce a introducir dos frecuencias de plasma diferentes y dependientes de la temperatura, $\omega_{pn}(T)$ y $\omega_{ps}(T)$. Por tanto, en el caso superconductor la respuesta dieléctrica tiene la forma

$$\varepsilon_s(\omega) = \varepsilon_\infty + \frac{i\pi(\omega_{ps}(T))^2}{2\omega}\delta(\omega) - \frac{\omega_{ps}^2(T)}{\omega^2} - \frac{\omega_{pn}^2(T)}{\omega^2 + i\gamma_0\omega} - \frac{\Omega_{mir}^2}{\omega^2 - \omega_{mir}^2 + i\Gamma_{mir}\omega}. \quad (14)$$

Cabe señalar que en el caso superconductor los portadores de carga son pares de Cooper formados por electrones acoplados y que presentan en conjunto un comportamiento estadístico de tipo bosónico.

En la figura 2 presentamos los resultados para el calor transmitido, Q , para una configuración de placas separadas por una distancia $L = 50\text{nm}$, suponiendo que una de ellas se mantiene una temperatura fija $T_1 = 120\text{K}$, mientras que la temperatura de la segunda varía entre $T = 10\text{K}$ y $T = 120\text{K}$. Observamos que, para $T > T_c$, el flujo de calor se incrementa a medida que la segunda placa disminuye su temperatura; sin embargo, justo en $T = T_c$, ocurre una disminución abrupta y el flujo de calor decrece en forma paulatina para temperaturas menores que T_c . En la figura 3 se presenta el equivalente al calor específico del sistema, representado por la derivada dQ/dT . Vemos que éste presenta una discontinuidad en T_c , en forma similar a la transición de fase termodinámica que sufren los materiales superconductores al pasar del estado normal al superconductor. Los resultados anteriores conducen a la especulación de diseñar nanomotores en donde parte del ciclo termodinámico esté asociado a esta transición.

FUERZAS DE CASIMIR

Según encontramos anteriormente, existe una densidad de energía asociada a las fluctuaciones cuánticas la cual, multiplicada por la densidad de modos del sistema tiene la forma

$$\rho_0(\omega) = \frac{\hbar\omega^3}{\pi^2c^3}. \quad (15)$$

Puede demostrarse que esta expresión es invariante relativista, de modo que las fluctuaciones de vacío no representan sistemas de referencia privilegiados y son entonces indetectables por observadores en un estado de movimiento inercial. Sin embargo, un observador en un sistema de referencia privilegiado (no inercial) puede detectar efectos asociados a las fluctuaciones cuánticas. En otras palabras, si la invariancia de Lorentz de un sistema se rompe por la presencia de fronteras materiales, campos externos, o movimientos acelerados, por ejemplo, entonces la densidad espectral de energía ya no

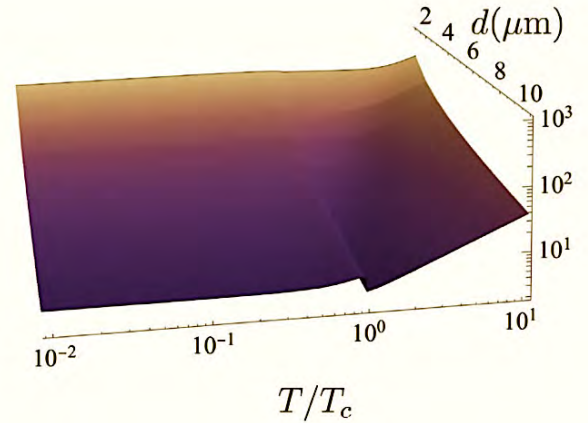


FIG. 4: Fuerza de Casimir (multiplicada por la distancia) como función de la temperatura reducida T/T_c y de la separación entre las placas, denotada aquí por d . Obsérvese la discontinuidad de la fuerza en $T = T_c$.

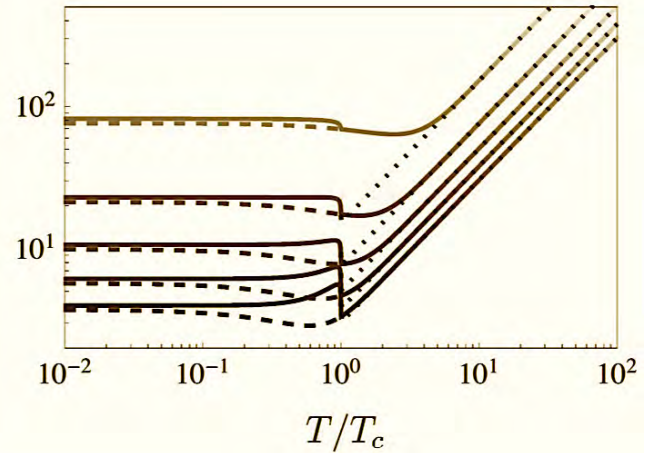


FIG. 5: Cortes de la superficie presentada en la figura 4 para la fuerza de Casimir (multiplicada por la distancia) como función de la temperatura reducida T/T_c para diferentes separaciones entre las placas. De la curva superior a la inferior, $d = 2, 4, 6, 8, 10\text{ nm}$.

corresponderá a la del sistema isotrópico y homogéneo implícitamente descrito por la Ec.(15).

El concepto de energía de punto cero del estado fundamental de un sistema electromagnético puede extenderse a una variedad de sistemas físicos de distinta índole. El comportamiento de sistemas colectivos constituidos por una gran cantidad de partículas cargadas en interacción mutua y con campos externos puede entenderse en términos de las oscilaciones de diferentes modos normales que permiten el transporte de energía, momento lineal y momento angular a través del sistema.

A nivel cuántico, dichas oscilaciones pueden ser descritas mediante conjuntos de osciladores armónicos cuyos niveles energéticos de estados excitados pueden asociarse a cuasipartículas cuánticas tales como los fonones (modos vibratorios en una estructura cristalina), plasmones (excitaciones coherentes de un plasma electrónico), excitones (excitaciones de un electrón y un agujero de carga), magnones (excitaciones coherentes de espín), etc. El modo fundamental de cada uno de estos sistemas, asociado a cero cuantos, determina el estado de vacío cuántico correspondiente. Las fluctuaciones de punto cero se manifiestan en una variedad de fenómenos tales como las fuerzas intermoleculares de Van der Waals, la emisión atómica espontánea, a la anchura fundamental de los niveles atómicos y, posiblemente, a la energía oscura del Universo.

Una de las más simples del vacío cuántico que ha permitido estudiar su estructura con mayor profundidad, es la de las fuerzas de Casimir. Éstas se originan en las fluctuaciones del campo electromagnético de vacío confinado por dos placas paralelas perfectamente conductoras y separadas por una distancia L . Un cálculo relativamente directo muestra que entre ellas debe aparecer una fuerza atractiva, con una magnitud

$$\frac{F}{A} = -\frac{\pi^2 \hbar c}{240 L^4} \quad (16)$$

en donde A es el área de las placas. Esta fuerza es muy pequeña a escalas macroscópicas y sólo es relevante a escalas nano- o micrométricas. Como consecuencia, sólo fue hasta años recientes en que los estudios experimentales de las fuerzas de Casimir alcanzaron la precisión necesaria para verificar con detalle las predicciones teóricas. Sin embargo, dichas predicciones no están basadas en la propuesta original de Casimir, sino en una teoría mucho más realista desarrollada posteriormente por Lifshitz, basada en el cálculo del tensor de esfuerzos del campo electromagnético y donde juega un papel fundamental el mecanismo de fluctuación-disipación, lo cual permite tomar en cuenta las propiedades dispersivas y disipativas de los materiales. La fórmula de Lifshitz para la fuerza por unidad de área se expresa normalmente en el espacio de frecuencias imaginarias

$$F(L, T) = -\frac{k_B T}{\pi} \sum_{\ell=0}^{\infty} \int_0^{\infty} d\beta \beta \kappa(i\zeta_{\ell}, \beta) \quad (17)$$

$$\times \sum_{\alpha=s,p} \left\{ \frac{\exp[2\kappa(i\zeta_{\ell}, \beta)L]}{r_{\alpha}^2(i\zeta_{\ell}, \beta)} - 1 \right\}^{-1},$$

en donde la suma se realiza sobre frecuencias de Matsubara $i\zeta_{\ell} = i(2\pi k_B T/\hbar)\ell$, con $\ell = 0, 1, 2, \dots$ y la prima en el símbolo de la sumatoria implica que el término con $\ell = 0$ debe multiplicarse por $1/2$. Las variables y funciones involucradas en esta expresión tienen la

misma interpretación que las empleadas en el caso de transferencia de calor. Por otro lado, la suma sobre frecuencias de Matsubara es, mediante una transformación analítica, equivalente a la integral sobre las frecuencias tomando en cuenta las contribuciones térmicas asociadas a la distribución de Planck. El uso de este tipo de formalismo representa una gran conveniencia en la realización de los cálculos, ya que en vez de involucrar términos fuertemente oscilatorios ocurren términos con decaimiento exponencial.

Similarmente al estudio realizado para la transferencia radiativa de calor, investigamos las fuerzas de Casimir entre placas superconductoras de YBCO separadas por una distancia L . En la figura 4 se presentan los resultados para la fuerza. Es claro que ocurre un incremento notable y abrupto en la fuerza en $T = T_c$, mientras que como función de la separación la fuerza presenta un comportamiento monótono y crece conforme disminuye la separación. La figura 5 representa diferentes cortes de la superficie presentada en la figura 4 como función de la temperatura reducida T/T_c para diferentes separaciones entre las placas. Puede observarse que la magnitud del salto relativo de la fuerza se incrementa para separaciones mayores, aunque en este caso las fuerzas son muy débiles y su medición representa un reto experimental.

En la actualidad ha sido posible construir dispositivos constituidos por micro-viguetas recubiertas por capas superconductoras y acopladas a cavidades con frecuencias de resonancia definidas capaces de detectar las variaciones de las fuerzas de Casimir en el tránsito del estado normal al superconductor. En el caso de superconductores normales, como el aluminio con $T_c \approx 1$ K, las variaciones han resultado indiscernibles. En el caso de los superconductores de alta T_c , de acuerdo a nuestras predicciones, esperaríamos que dicha detección fuese factible. Esperamos que en un futuro próximo las investigaciones experimentales puedan abocarse a este tipo de materiales.

Para concluir, cabe mencionar que se han podido medir las fuerzas tipo Casimir inducidas por el campo de punto cero de fonones en sistemas acústicos. Asimismo, se han calculado las fuerzas tipo Casimir asociadas a fluctuaciones cuánticas de electrones de un material, o a fluctuaciones del campo de Higgs en espacio-tiempos con dimensiones compactificadas. En general, podemos concebir las fuerzas de Casimir asociadas a cualquier tipo de excitación cuántica.

REFERENCIAS

Fuerzas del vacío cuántico. Carlos Villarreal, PP 381-386. Publicado en *Analogías y Conexiones de la Física*, Ed. por Karen Volke, Jorge Flores, Octavio Miramontes, Sergio Snchez CopIt-ArXives (2020), ISBN: 978-1-938128-19-6 ebook

The Quantum Vacuum. An Introduction to Quantum Electrodynamics Peter W. Milonni. Academic Press Inc. Boston (1994).

Near-field radiative heat transfer and noncontact friction. A. I. Volokitin and B. N. J. Persson, Reviews of Modern Physics, Vol 79. 1291-1329 (2007).

Near-field radiative heat transfer between high-temperature superconductors. Shunashi Guadalupe

Castillo López, Giuseppe Pirruccio, Carlos Villarreal, and Raúl Esquivel-Sirvent. Scientific Reports 10, 16066 (2020), doi.org/10.1038/s41598-020-73017-z .

Role of Relaxation Rates in the Casimir Force between High-Tc Superconductors. Shunashi Guadalupe Castillo López, Carlos Villarreal, Giuseppe Pirruccio, and Raúl Esquivel-Sirvent. Universe 7, 69 (2021). <https://doi.org/10.3390/universe7030069> .

Casimir forces and high-Tc superconductors. Carlos Villarreal and Santiago F. Caballero Benítez. Physical Review A 100, 042504 (2019). doi:10.1103/PhysRevA.100.042504

Vacuum Fluctuations and Symmetry Breaking in Spaces with Nontrivial Topology. C. Villarreal. Physical Review D 51, 2959 (1995).

Levitadores acústicos uniaxiales: explorando sus límites para mejorar su desempeño

Victor Contreras, Instituto de Ciencias Físicas, UNAM Campus Morelos.

1. Levitación acústica

La levitación acústica se refiere a la habilidad de atrapar y sostener con ondas ultrasónicas a objetos micrométricos y milimétricos en fluidos como el aire o el agua. A diferencia de otras técnicas, la levitación acústica no está restringida a las propiedades físicas de los objetos como su respuesta eléctrica o magnética. Sin embargo, para objetos pequeños comparados con la longitud de onda ultrasónica, la fuerza responsable de levitar es proporcional a la densidad de los objetos [1]. Esta cantidad física representa una de las limitaciones en el desempeño de los levitadores acústicos en situaciones prácticas.

Los levitadores acústicos se desarrollaron en los años 70's para producir ambientes de microgravedad en investigaciones realizadas en el Jet Propulsion Laboratory de la NASA [2]. A lo largo de las siguientes cinco décadas esta tecnología ha evolucionado y ha demostrado su utilidad en diversas aplicaciones. Por ejemplo, la levitación acústica ha hecho posible la suspensión y manipulación de sustancias para producir reacciones químicas sin el uso de contenedores que puedan contaminar o interferir en la reacción [3]; el confinamiento y manipulación "sin contacto" de especies biológicas para su posterior estudio [4]; el estudio de propiedades físicas y químicas de líquidos [5]; el desarrollo de pantallas 3D basadas en la holografía acústica [6], etc.

2. Levitación acústica por medio de ondas estacionarias

Los modos de vibración asociados a un sistema resonante presentan un patrón de oscilación característico conocido como onda estacionaria. La generación de ondas estacionarias dentro de cavidades uniaxiales resonantes es uno de los métodos de levitación acústica más comunes y explorados debido a su potencial científico, tecnológico e industrial. La onda estacionaria ultrasónica es el resultado de la interferencia de las ondas acústicas contra propagantes y sus múltiples reflexiones dentro de la cavidad, y están asociadas a las distribuciones o gradientes de presión y a las fuerzas responsables de levitar objetos en posiciones específicas de la trampa acústica. Un levitador uniaxial de cavidad resonante está formado por un emisor y un reflector (el reflector puede sustituirse por otro emisor que opere a la misma frecuencia) donde las superficies que emiten y reflejan las ondas acústicas están separadas una distancia que define la longitud de la cavidad. La figura 1 (a) muestra esquemáticamente un levitador acústico que está representado por el sólido de revolución

que se obtiene al girar 180° al plano respecto de la línea punteada roja que representa el eje de simetría.

En el caso de objetos pequeños comparados con la longitud de onda de la onda acústica, la levitación acústica ocurre en los nodos de la onda estacionaria. Si las superficies de la cavidad son planas, la onda estacionaria reside dentro del volumen cilíndrico limitado por dichas superficies como lo ilustra la Figura 1(b). Para formar la cavidad resonante, la longitud de la cavidad debe ajustarse a un tamaño múltiplo de media longitud de onda, λ , asociada a la vibración del emisor ultrasónico:

$$L = n(\lambda/2) \quad (1).$$

Para una cavidad con superficies cóncavas (Fig. 1(c)), el campo ultrasónico se enfoca alrededor del centro de la curvatura y la onda estacionaria está contenida en el volumen de un hiperboloide. Esta distribución concentra la presión radialmente y aumenta considerablemente su amplitud en los antinodos centrales y, consecuentemente, aumenta la estabilidad espacial de los objetos levitados. En el caso de superficies cóncavas, la longitud de la cavidad no es un múltiplo entero de media longitud de onda debido a que la condición de resonancia depende de la curvatura de las superficies que conforman la cavidad. Sin embargo, si los radios de curvatura son considerablemente grandes comparados con λ , la longitud de resonancia de la cavidad se puede aproximar a la ecuación (1): $L \approx n(\lambda/2)$.

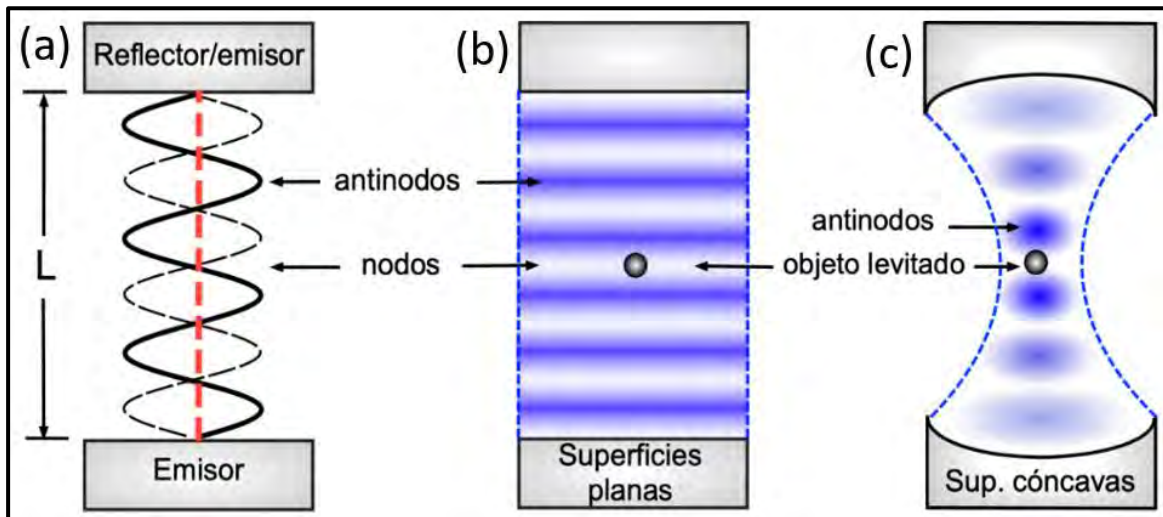


Figura 1. Representación esquemática de levitadores acústicos uniaxiales de cavidad resonante. (a) La cavidad resonante está constituida por superficies emisoras y reflectoras separadas una distancia múltiplo de media longitud de onda. (b) Onda estacionaria de un sistema de levitación de superficies planas en modo resonante $n = 6$. (c) En el caso de un levitador con superficies cóncavas, la onda estacionaria se enfoca en el centro de la cavidad ($n = 6$).

3. Levitadores acústicos uniaxiales

En la literatura científica podemos encontrar información acerca de cavidades acústicas y sistemas de levitación basados principalmente en transductores de alta potencia que requieren entre los 100 y 1000 V para operar. Sin embargo, en la última década se ha desarrollado otro tipo de levitadores basados en una tecnología de transductores más simple y compacta. El uso de ambas tecnologías para el diseño y desarrollo de levitadores acústicos se discute brevemente en esta sección.

3.1 Levitadores basados en transductores tipo Langevin

La generación continua de ondas ultrasónicas se basa principalmente en el uso de transductores que consisten en discos de material piezoeléctrico apilados entre electrodos metálicos y sujetos a masas metálicas por medio un tornillo como se ilustra en la Figura 2 (a). Cuando un voltaje periódico se aplica a los discos piezoeléctricos, éstos se deforman de manera que producen un movimiento en la dirección vertical que es proporcional a la amplitud del voltaje aplicado. Estos dispositivos se conocen como transductores Langevin y se utilizan en aplicaciones diversas que abarcan limpieza ultrasónica, soldadura de plásticos y metales, maquinado de precisión, dispositivos médicos y dentales, levitación acústica, etc. En la mayoría de las aplicaciones, el transductor se acopla a un amplificador mecánico para lograr la amplitud requerida.

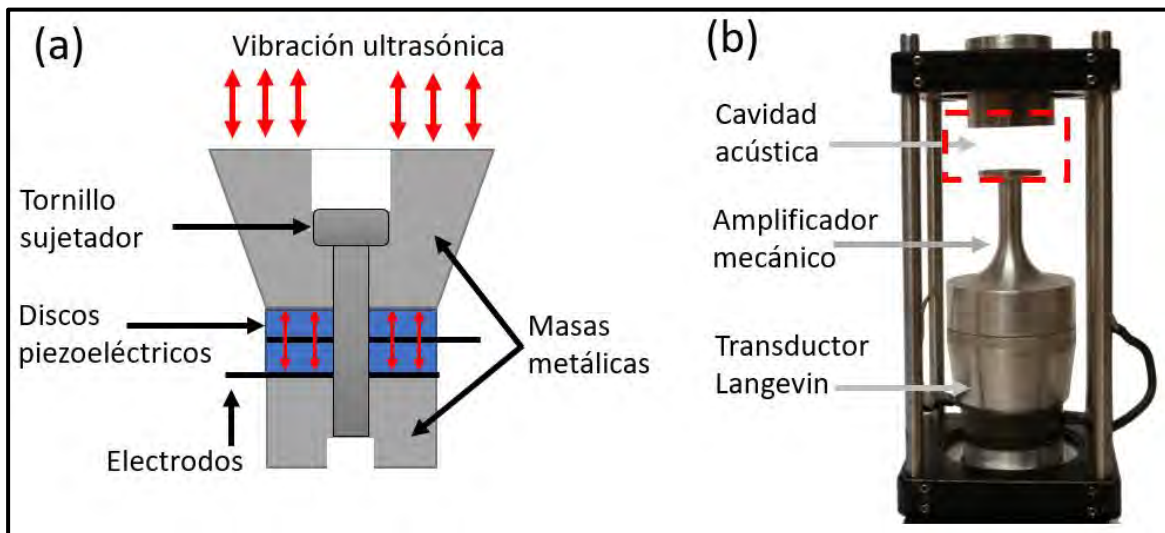


Figura 2. (a) Ilustración de un corte transversal de un transductor Langevin y sus componentes. (b) Levitador acústico basado en un transductor tipo Langevin, acoplado a un amplificador mecánico para incrementar la amplitud de la onda ultrasónica. La cavidad acústica se forma al contraponer una superficie reflectora a la superficie del amplificador mecánico.

En el caso de la levitación acústica, generalmente se han diseñado levitadores capaces de levantar esferas de poliestireno y muestras líquidas. Sin embargo, en la literatura científica

podemos estudios paramétricos orientados a diseñar de cavidades acústicas uniaxiales con la capacidad de levitar objetos de alta densidad (Figura 2(b)). Por ejemplo, en 1980 Oran et. al reportaron un levitador compuesto por un transductor y un reflector (T-R) de superficies cóncavas con la capacidad de levitar muestras de platino ($\rho \sim 20 \text{ g/cm}^3$) al operar en condiciones de resonancia [7]. En el año 2002, Wei et al. demostraron que, posterior a la optimización de parámetros geométricos como los radios de curvatura de las superficies cóncavas y la longitud de una cavidad T-R, es posible levitar objetos sólidos y líquidos de cualquier densidad al levitar esferas milimétricas de iridio y mercurio [8]. Sin embargo, para producir las fuerzas capaces de levitar tales materiales, el levitador debe operar en el primer modo resonante ($n = 1$). Esta restricción reduce la longitud de la cavidad a pocos milímetros y limita el campo de visión en la dirección axial. Un campo visual reducido limita el acceso óptico a las muestras levitadas y, consecuentemente, a las aplicaciones prácticas del dispositivo. Esta limitación es común en levitadores basados en transductores tipo Langevin, ya que, para alcanzar un buen desempeño, generalmente las trampas se diseñan para operar en sus primeros modos resonantes. La figura 2 muestra un levitador acústico T-R construido a partir de un transductor tipo Langevin, que opera a 40 KHz, y un reflector de superficie cóncava. La longitud de la cavidad acústica es de aproximadamente 2.5λ .

3.2 Levitadores basados en arreglos de transductores en fase

En la última década se ha desarrollado un novedoso tipo de levitadores con características atractivas. En estos levitadores, los transductores tipo Langevin son sustituidos por arreglos de múltiples transductores más compactos y eficientes. El uso de transductores más compactos permite ordenarlos en arreglos o matrices que operan en fase de manera que la superficie emisora del levitador esté discretizada y vibre de manera similar a la superficie continua de un transductor tipo Langevin. Esta discretización tiene dos ventajas principales con respecto a los levitadores basados en un solo transductor: 1) se pueden diseñar superficies emisoras de mayor área y de geometrías más variadas de manera simple; y 2) se pueden manipular los transductores independientemente y controlar electrónicamente su fase, abriendo la posibilidad de modular el campo acústico producido por los arreglos de transductores. Otra característica atractiva es que, al ser transductores más eficientes, es posible generar amplitudes de presión similares a los transductores tipo Langevin, pero con un consumo menor de potencia eléctrica. Esta característica es importante para el desarrollo de levitadores más compactos y robustos y para la miniaturización de instrumentación científica y tecnológica.

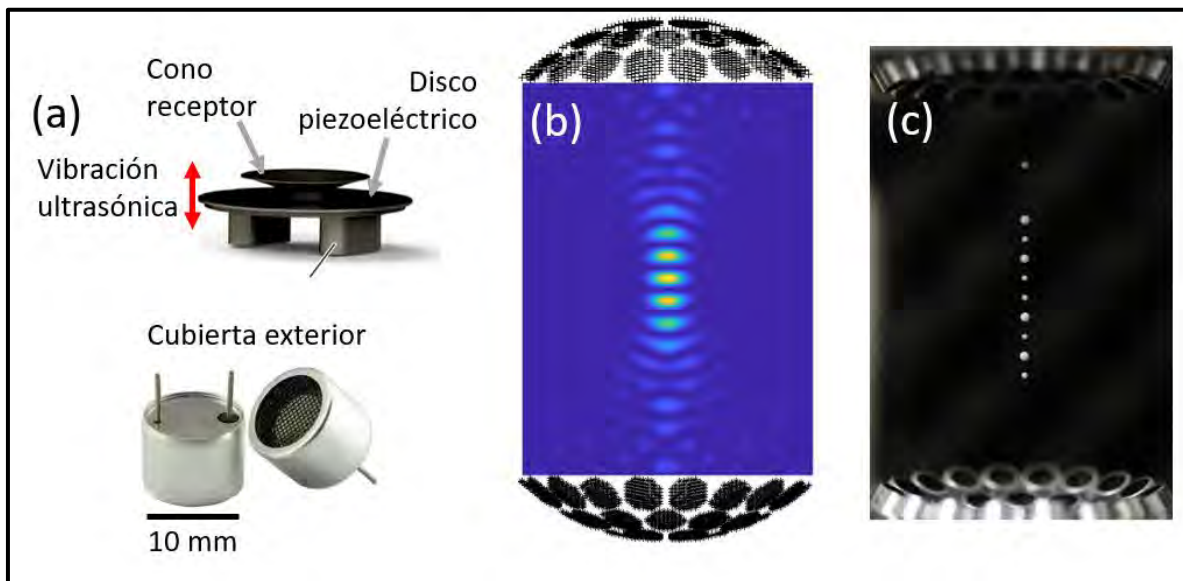


Figura 3. (a) Diseño de los sensores ultrasónicos de “estructura abierta” utilizados en la nueva generación de levitadores acústicos. (b) Simulaciones numéricas de la distribución de presión acústica dentro de una cavidad resonante construida a partir de arreglos de transductores compactos. (c) Imagen del sistema de levitación con partículas de poliestireno en distintos nodos de la trampa acústica. La longitud de la cavidad es de aproximadamente 13λ .

La figura 3 (a) muestra la estructura de los transductores compactos y consta de un disco piezoeléctrico acoplado a un disco metálico y a una estructura emisora/receptora en forma de cono para radiar de manera más eficiente al aire las ondas ultrasónicas generadas por la oscilación de los discos. El cono, a su vez, acopla de manera efectiva ondas ultrasónicas que recibe y las concentra en la parte central del disco, por lo que el dispositivo funciona como emisor y receptor. Al funcionar como emisor-receptor (transceptor), estos dispositivos se conocen como sensores ultrasónicos.

Algunos estudios recientes han demostrado que las trampas acústicas de arreglos en fase son capaces de levitar objetos de alta densidad. En 2017 Marzo et al., reportaron un levitador acústico basado en un diseño de cavidad concéntrica uniaxial al que llamaron TinyLev [9]. Los autores describen el diseño y la caracterización experimental del TinyLev, capaz de levitar objetos con densidades de hasta 6.5 g/cm^3 y, de acuerdo con sus resultados numéricos, sugieren la posibilidad de mejorar el desempeño del levitador al reducir la curvatura de sus superficies cóncavas. En 2021, Contreras y Marzo [10] reportaron el diseño y el desempeño de un levitador similar al TinyLev y mostrando que, al ajustar finamente la longitud de la cavidad y operar al levitador en modo resonante, es posible levitar esferas de acero y mercurio, demostrando una mejora considerable en el desempeño de levitación de las cavidades acústicas basadas en arreglos en fase.

4. Investigaciones actuales en el ICF-UNAM

Actualmente (2019 -2021) en el Laboratorio de Óptica del ICF-UNAM se realizan estudios paramétricos de levitadores uniaxiales basados en arreglos de transductores en fase con el objetivo de optimizar su desempeño. La parametrización busca maximizar numéricamente la amplitud del “exceso de presión promedio” en los antinodos centrales en función de parámetros específicos de las cavidades ultrasónicas. Una vez optimizado numéricamente el diseño de la cavidad, se construyen los levitadores y se realizan las pruebas experimentales correspondientes. Los parámetros estudiados numéricamente incluyen: la longitud de la cavidad, la curvatura del arreglo de transductores, la geometría de la cavidad y la fase relativa entre ambos arreglos de transductores. La cavidad acústica se optimiza experimentalmente utilizando técnicas ópticas como la deflectometría schlieren, que permite visualizar los gradientes de densidad del medio y asociarlos a la distribución de presión dentro de los levitadores [10]. La Figura 4 muestra un par levitadores acústicos de arreglos de transductores diseñados y construidos en el laboratorio de Optica del ICF-UNAM y que tienen la capacidad de levitar objetos de alta densidad.



Figura 4. (a) Levitador acústico T-T basado en dos arreglos de 18 transductores levitando una esfera de acero de 2 mm de diámetro. El diseño es una versión miniaturizada del levitador TinyLev. En la fotografía se observa el gradiente de presión asociado a la onda estacionaria a través de deflectometría schlieren. (b) Levitador acústico T-R basado en un arreglo de 8 transductores y un reflector cóncavo levitando una gota de agua.

5. Agradecimientos

Agradecimientos a los Programas de Apoyo a Proyectos de Investigación e Innovación Tecnológica (PAPIIT) IA101319 e IN109221 de DGAPA-UNAM por brindar los recursos para la adquisición del equipo y material utilizado en el trabajo descrito en este documento.

6. Referencias

- [1] M. A. B. Andrade, N. Pérez, and J. C. Adamowski, "Review of Progress in Acoustic Levitation," *Brazilian J. Phys.*, vol. 48, no. 2, pp. 190–213, 2018.
- [2] T. WANG, M. SAFFREN, and D. ELLEMAN, "Acoustic chamber for weightless positioning," pp. 0–5, 1974.
- [3] S. V. Ushakov, J. Niessen, D. G. Quirinale, R. Prieler, A. Navrotsky, and R. Telle, "Measurements of density of liquid oxides with an aero-acoustic levitator," *Materials (Basel)*, vol. 14, no. 4, pp. 1–13, 2021.
- [4] Z. Yang *et al.*, "Light sheet microscopy with acoustic sample confinement," *Nat. Commun.*, pp. 1–8, 2019.
- [5] V. Contreras, R. Valencia, J. Peralta, H. Sobral, M. A. Meneses-Nava, and H. Martinez, "Chemical elemental analysis of single acoustic-levitated water droplets by laser-induced breakdown spectroscopy," *Opt. Lett.*, vol. 43, no. 10, p. 2260, 2018.
- [6] R. Hirayama, D. M. Plasencia, N. Masuda, and S. Subramanian, "A volumetric display for visual, tactile and audio presentation using acoustic trapping," *Nature*, vol. 575, no. November, 2019.
- [7] W. A. Oran, L. H. Berge, and H. W. Parker, "Parametric study of an acoustic levitation system," *Rev. Sci. Instrum.*, vol. 51, no. 5, pp. 626–631, 1980.
- [8] W. J. Xie, C. D. Cao, Y. J. Lü, and B. Wei, "Levitation of Iridium and Liquid Mercury by Ultrasound," *Phys. Rev. Lett.*, vol. 89, no. 10, pp. 1–4, 2002.
- [9] A. Marzo, A. Barnes, and B. W. Drinkwater, "TinyLev: A multi-emitter single-axis acoustic levitator," *Rev. Sci. Instrum.*, vol. 88, no. 8, 2017.
- [10] V. Contreras and A. Marzo, "Adjusting single-axis acoustic levitators in real time using rainbow schlieren deflectometry," *Rev. Sci. Instrum.*, vol. 92, no. 1, 2021.

Galaxy - dark matter connection in Cosmology Conexión galaxia - materia oscura en cosmología

Sébastien Fromenteau

December 11, 2021

Abstract

Cosmology is the science that study the evolution of the Universe comparing observations (mostly given by astrophysics objects) with theoretical model that embed General Relativity, Particle physics and Large Scale Structure (LSS) models. In particular, the LSS is governed by the so called Dark Matter that dominates the total matter component of the Universe. The standard matter we can interact with, abusively called 'baryonic matter', represent only 15 % of the matter content. However, the baryonic matter is the one we can observe. It means that cosmology use partial and biased tracers of the underlying total matter distribution, using galaxies, intergalactic medium or galaxy clusters. We will see that the bias is of first importance and that can be decomposed in two parts: halo bias and galaxy-halo connection. In order to understand well the propose of the connection bias we have first to dig a bit on the halo bias.

1 Introduction

1.1 Information compression and statistics

The way we can compress the information from spatial distribution is to use the spatial N-point statistics. In order to understand this we can do an analogy with the way we can characterize a probability distribution function (PDF) $p(x)$ with its moments, where the first moment μ_1 corresponds to the mean :

$$\mu_1 = \int_{-\infty}^{+\infty} p(x) \times x dx, \quad (1)$$

the second moment μ_2 corresponds to the variance:

$$\mu_2 = \int_{-\infty}^{+\infty} p(x) \times (x - \mu_1)^2 dx, \quad (2)$$

and the n -th moment is defined by:

$$\mu_n = \int_{-\infty}^{+\infty} p(x) \times (x - \mu_1)^n dx. \quad (3)$$

If we know all the value of all the moments then we can perfectly draw the PDF. However you should ask : 'why is it interesting? Know the value of $p(x)$ for all x is the

same than to know all the moments.' Indeed, if you really need all the moments you are not compressing any information. But in general you just need the first moments to characterize pretty well your PDF. In particular, the limit-central theorem implies that the sum of complex phenomenon will tend to follow a normal distribution (i.e. a Gaussian). And the formula of a Gaussian is given by:

$$\mathcal{N}(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x - \mu)^2}{2\sigma^2}\right), \quad (4)$$

where $\mu = \mu_1$ is the mean and $\sigma^2 = \mu_2$ is the variance. So if you know that process your are studying is supposed to be Gaussian or very close to be, you can just evaluate the mean and the variance with your data and you have your distribution. That's the basic idea of statistics! Accurately evaluate the underlying probability distribution with a finite number of data.

Because we have finite data, it is not possible to use the continuous definitions Eq.1-3 to evaluate the moments. We need to use estimators and you will see that you know very well the first two. In the case I have access to N data points X_i $i = 1, 2, 3...N$, I can estimate the mean of the distribution $\langle X \rangle = \hat{\mu}_1$ as follow:

$$\hat{\mu}_1 = \langle X \rangle = \frac{1}{N} \sum_{i=1}^N X_i, \quad (5)$$

where $\hat{\mu}_1$ means the estimator of the first moment. You probably probably already used it thousands of times. As you know, this estimator works well but also have an error due to the finite number of data. The mean estimator reads as the sum of N values X_i and then $\langle X \rangle$ is supposed to follow a Gaussian distribution due to limit-central theorem. For example I can generate $N = 1000$ data following the uniform distribution $X_i \in [0, 1]$ and then evaluate the mean using the estimator $\langle X \rangle$ as presented in Fig 1. We can see that the mean estimation is pretty good but not perfect (0.5058 rather than 0.5). However we can see on the right panel of Fig 1 that doing the same experience 10,000 times and estimating the mean of the means we find a value very close to the correct one. In statistics it corresponds to a non-biased estimator.

In fact we can easily interpret this estimator knowing that $1/N$ is the frequentist probability associated to

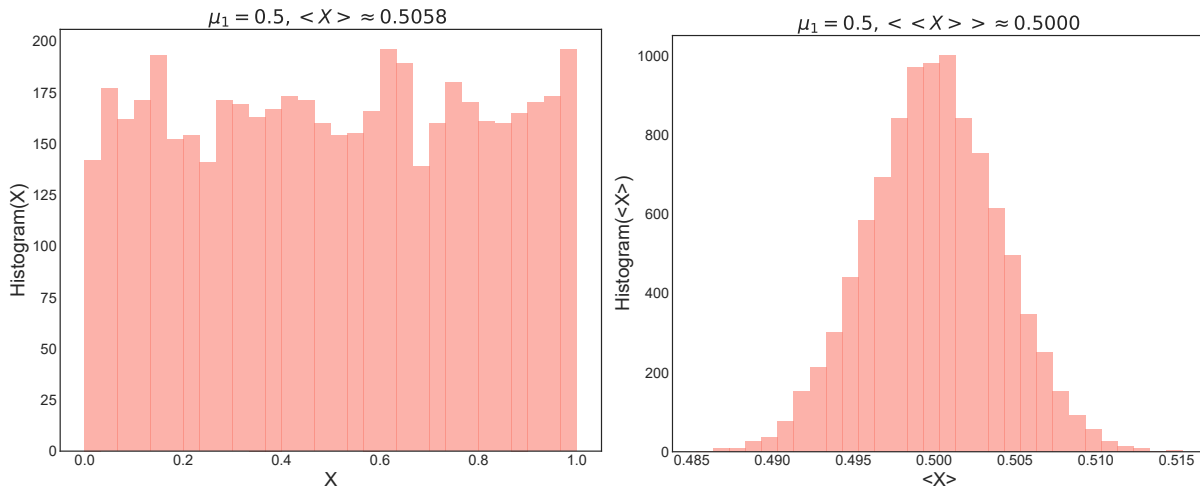


Figure 1: **Left panel:** histogram of the X_i following uniform distribution $\mathcal{U}(0, 1)$ for $N = 5000$. The mean value encountered is $\langle X \rangle 0.5058$ that is close to the real value $\mu_1 = 0.5$ but not exact. **Right panel:** histogram of the mean values encountered doing 10,000 times the experience presented in the left panel. We can see that the distribution is a Gaussian as expected by the central-limit theorem. We can see the width of the Gaussian that provide the information of the variance of the estimator when the difference between the two lines corresponds to the bias of the estimator. Here we can see that the mean estimator $\langle X \rangle$ is not biased.

each data point. So we wrote the discrete version of μ_1 with $\langle X \rangle = \sum_{i=1}^N p(X_i)X_i$ and we should extend this intuitive way to create the variance estimator as:

~~$$\hat{\mu}_2 = Var(X) = \frac{1}{N} \sum_{i=1}^N (X_i - \langle X \rangle)^2, \quad (6)$$~~

but one can demonstrate that this is a biased estimation of the second moment. In the left panel of Fig. 2 is presented the variance estimation distribution for 10,000 experiences using data number $N = 10$. We choose to reduce N to make the bias evident to see. Indeed, we can observe that the mean value of the 10,000 experiences give a value of 0.075 while the expected one for the uniform distribution is $\mu_2 = 1/12 = 0.083333..!$ One can use the Jackknife sampling method to show that this estimator is biased by a factor of $(N - 1)/N$ and so correct estimator is given by:

$$\hat{\mu}_2 = Var(X) = \frac{1}{N-1} \sum_{i=1}^N (X_i - \langle X \rangle)^2, \quad (7)$$

The higher order moments have different estimators that have advantages and drawbacks. We start to enter in the problematic of the estimators, deal with the bias and variance of each one. To do science, lower the bias and variance of each one. To do science, lower the bias better is, however the less biased estimators generally come with the higher variance. So the error associated with your estimation is large but you know that if you can repeat the experience a large number of times, the mean will converge to the correct value. Sometimes, you should prefer to use a balanced estimator that provide to you the best compromised result: the estimator that minimize the variance giving a bias negligible compare to the intrinsic error of the measurement.

1.2 2-pt correlation function in a nutshell

We have introduced important basic concepts, using an intuitive example with the PDF, that will help to understand concept of n -pt correlation function.

A way to characterize a random process is to estimate the correlation functions (different orders will correspond to the number of points we use). The simplest and more important is the 2-pt correlation function. If we have a random distribution of points with a mean density $\bar{n} = N/V$; where N is the number of points in the total volume V ; the number of points we can expect to see in a small volume dV is given by: $dN = \bar{n}dV$. If we consider two different small volume dV_1 and dV_2 we expect to observe in average $dN_1 = \bar{n}dV_1$ and $dN_2 = \bar{n}dV_2$ points respectively. So, we expect to get an average number of pairs of points between the two volumes equal to:

$$dP_{1,2} = dN_1 \times dN_2 = \bar{n}^2 dV_1 dV_2. \quad (8)$$

Because we generally use only the counting around existing points, we are interested in the number of pairs existing between this point and the points inside a small volume dV . It directly corresponds to the number of expected points in the later volume:

$$dP = 1 \times dN = \bar{n}dV. \quad (9)$$

If the distribution is not exactly random in positions, we expect to obtain a different quantity. So we let the possibility to have an excess or default respect to the randomly expected value writing:

$$dP(\vec{r}) = 1 \times dN = \bar{n}dV[1 + \xi(\vec{r})], \quad (10)$$

where $\xi(\vec{r})$ is the 2-pt correlation function. If the distribution of the points is random, the number of pairs will

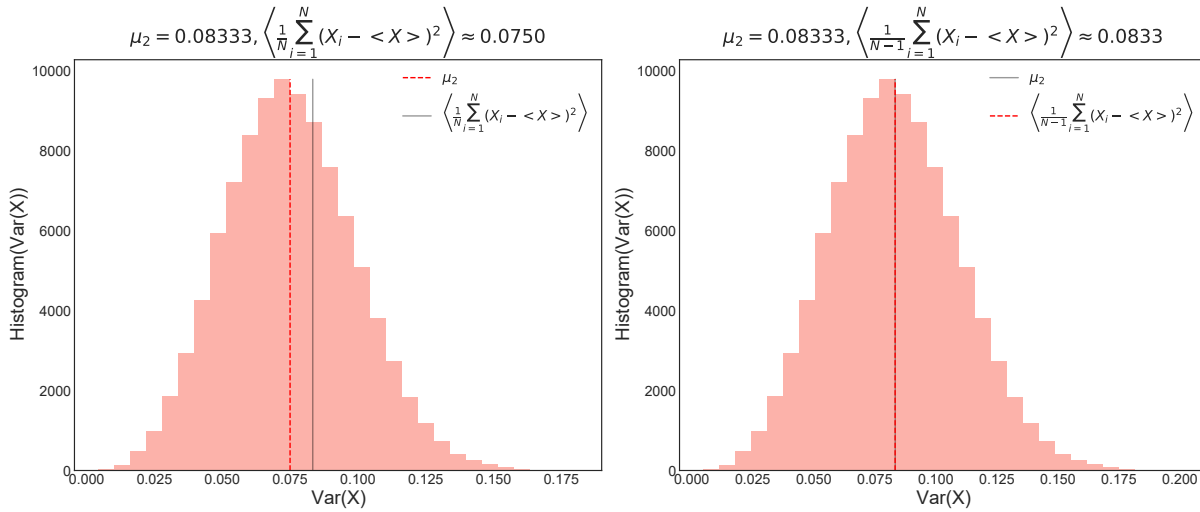


Figure 2: **Left panel:** histogram of the uncorrected variance estimation for X_i following uniform distribution $\mathcal{U}(0, 1)$ for $N = 10$. The mean value for this variance estimator is different of the expected value (0.075 while the expected value is $\mu_2 = 1/12 = 0.083333\dots$). So this estimator of the variance is biased by a factor $(N - 1)/N$ as it is possible to demonstrate using Jackknife sampling method. **Right panel:** histogram of the correct variance estimation. We can see that this estimator is non-biased and so the one one should use.

be compatible with $\bar{n}dV$ and so the correlation function will be null. In the other case, we will find excess and defaults in particular directions and orientations. If we consider an isotropic distribution (like the Universe if we believe in the Cosmological Principle) the deviation from the random expected number of pair have to be independent of the orientation and so will depend only on the distance $|\vec{r}|$. Moreover, in the case isotropy, we can directly consider the all shell over the point with radius r . So we can recast the equation using the volume in the shell as:

$$dP(r) = \bar{n} \times 4\pi r^2 dr [1 + \xi(r)], \quad (11)$$

which is the most common way to express the 2pt-correlation function in cosmology. In all the reasoning we done before we use the number of pair of points we expect to measure so the natural way to estimate the correlation function will be using the pair counts at each scale r and compare it with the expected value for a random distribution. In the case of a simple realization, we can estimate the correlation function as:

$$\hat{\xi}(r) = \frac{DD(r)}{N\bar{n} \times 4\pi r^2 dr} - 1, \quad (12)$$

where $DD(r)$ is the number of pairs we count at a distance $\in [r, r + dr]$ considering the N Data points. As simple case, we refer to a periodic box for which there is no limits in the pair counting. Indeed, we can always draw a complete shell around each point in the limit of the size of this box. However, the reality is different and the expected number of pairs for the random realization is in general impossible to evaluate theoretically. For this reason, we create a random catalog reproducing the geometry containing the data points and we can then compare the pair counts between the data $DD(r)$

and the random realization $RR(r)$ as:

$$\hat{\xi}(r) = \frac{N \times DD(r)}{N \times RR(r)} - 1 = \frac{DD(r)}{RR(r)} - 1. \quad (13)$$

Moreover, we in order to reduce the variance in the estimation of $\hat{\xi}(r)$, we can increase the density of the random sample. We also need to take in to account that we will measure more pairs in the random than in the data points. If we multiply the density by a factor β then we will have $\beta N(\beta N - 1) \approx \beta^2 N^2$ pairs when we will measure $N(N - 1) \approx N^2$ pairs for the data:

$$\hat{\xi}(r) = \beta^2 \frac{DD(r)}{RR(r)} - 1, \quad \beta \approx \frac{\bar{n}_{random}}{\bar{n}_{data}}. \quad (14)$$

This estimator is a good one if the number of data points is very large. However it exists different estimators to evaluate the 2pt-correlation function and the one optimizing the variance and the bias of the estimation is provided by [LS1993]:

$$\hat{\xi}_{L-S}(r) = \frac{DD(r) - 2DR(r) + RR(r)}{RR(r)}, \quad (15)$$

where $DR(r)$ is the number of pairs we can do between points from the random and the points from the data. It is possible to do it since we reproduce the geometry of the data in the random catalog. While the form looks very simple, the demonstration to show the efficiency of this estimator is pretty hard and developed in the reference. Find the best estimator of n -pt statistics is pretty hard, but the idea is the similar to the one discuss for the variance estimator in the previous section.

In case the distribution of the field does not follow a Gaussian process, the higher order statistics is need to characterize the distribution similarly to the

moments for non-gaussian PDF. The 3-pt correlation function analysis is actually used on the galaxy distribution in order to discriminate between different gravitation models as well as on the Cosmic Microwave Background temperature map in order to find primordial non-Gaussianities related to the inflationary period. Both cases use triples of galaxies or temperature (i.e. $DDD(r), DDR(r)...$) or the Fourier transform analysis named Bispectrum. This topic is of great interest but we will not develop more in this lecture. We want to focus on the bias introduced by the tracers on our estimations and the particular role of the galaxy - dark matter connection.

We introduced the estimators of 2pt-correlation function that we can apply on galaxy distribution observed in data. If we do that is because we are able to predict the value of the 2pt-correlation function of the matter perturbations. From the Inflation model we can predict an initial power spectrum (Fourier transform of the 2pt-correlation function) that provide to us the initial conditions of the matter perturbations $\delta_m(\vec{k})$ that are Fourier transform of the matter perturbations in real space $\delta_m(\vec{r})$. Then we can calculate the evolution of the individual perturbation modes on the linear perturbation limit, and combine them when start to be to large to neglect the second order terms (and higher orders). The last case is generally called 'Perturbation Theory' in cosmology and regroup a large range of approaches (Standard PT, Lagrangian PT, Effective PT, kinetic-PT...). We finally obtain the prediction for the power-spectrum and 2pt-correlation function for the matter perturbations (dominated by dark matter) at a given time for a given cosmological model.

The definition of the power-spectrum¹ is the variance of the Fourier space perturbations $\delta_m(\vec{k})$:

$$2\pi^3 P_{mm}(k) = \left\langle \delta_m(\vec{k}) \cdot \delta_m(\vec{k})^* \right\rangle_{|\vec{k}|=k} = \left\langle |\delta_m(\vec{k})|^2 \right\rangle_{|\vec{k}|=k}, \quad (16)$$

and the 2pt-correlation function as the excess/default of correlation at given separation scale r :

$$\xi_{mm}(r) = \langle \delta_m(\vec{x}) \cdot \delta_m(\vec{x} + \vec{r}) \rangle_{|\vec{r}|=r}, \quad (17)$$

that use the global matter information while we will estimate the correlation function of the galaxies. It should be the same if the galaxies were a perfect tracer of the whole matter distribution. Of course, galaxies do not trace perfectly the matter distribution, they live in the densest regions.

1.3 Tracer bias

So we will infer the galaxy correlation function of the galaxy density perturbation $\delta_g(\vec{r})$:

$$\xi_{gg}(r) = \langle \delta_g(\vec{x}) \cdot \delta_g(\vec{x} + \vec{r}) \rangle_{|\vec{r}|=r}. \quad (18)$$

¹we will just use the isotropic power spectrum for a real value field

Consider the galaxy as a tracer of the whole matter is equivalent to assume that it exist a function f :

$$f[\delta_m(\vec{r})] = \delta_g(\vec{r}). \quad (19)$$

Because this function depends on the value of the matter density field and the position it existence is not a problem. However, it is impossible to deal with this kind of definition depends on everything on the position and so have a number of degree of freedom as large as the number of galaxy used on the observation. So a first assumption one can make is the non-locality of the connection:

$$f[\delta_m] = \delta_g, \quad (20)$$

stipulating that the connection only depends on the density whatever the position \vec{r} . One can use a Taylor expansion to define in a general manner the connexion:

$$\delta_g = b_0 + b_1 \delta_m + \frac{b_2}{2} \delta_m^2 + \dots, \quad (21)$$

where the parameters b_i will refer to the bias terms. The first order (b_1) is known as the linear bias and was used in most of the studies of galaxy clustering since few years ago. The reason is the simplicity of the theoretical treatment and the single free nuisance parameter to add in the data analysis. It was totally justified to use it due to the statistical errors greater than the error implied by doing this simplification. The link between the observed 2pt-correlation function and power spectrum to the theoretical one read as:

$$\begin{aligned} \xi_{gg}(r) &= \langle b_1 \delta_m(\vec{x}) \cdot b_1 \delta_m(\vec{x} + \vec{r}) \rangle_{|\vec{r}|=r}, \\ \xi_{gg}(r) &= b_1^2 \langle \delta_m(\vec{x}) \cdot \delta_m(\vec{x} + \vec{r}) \rangle_{|\vec{r}|=r}, \\ \xi_{gg}(r) &= b_1^2 \xi_{mm}(r), \end{aligned} \quad (22)$$

and

$$P_{gg}(k) = b_1^2 P_{mm}(k). \quad (23)$$

We present the effect of linear bias on the 2pt-correlation function of the Luminous Red Galaxies from the CMASS sample in figure 4. However we reach the moment we can not continue with this approximation and we have to deal with more complex bias models. En particular we have to deal with scale dependence of the bias when we want to probe modification to the gravitational model or the impact of neutrinos on the large scale structure of the Universe.

We first need to understand the meaning of a tracer and then the associated bias it can produce on the estimators. The cosmological model stipulate that energy content of the Universe today is close to be 25% of dark matter (Cold Dark Matter to be more specific), 70% of Dark Energy, and 5% of normal matter and that the large structure formation is lead by the Dark Matter. An initial overdensity of matter will slowly increase by gravitational attraction up to reach a special value that correspond to the 'Turn-around' corresponding to a collapse. What happened at this time? The overdensity had a local curvature that start to win over the local

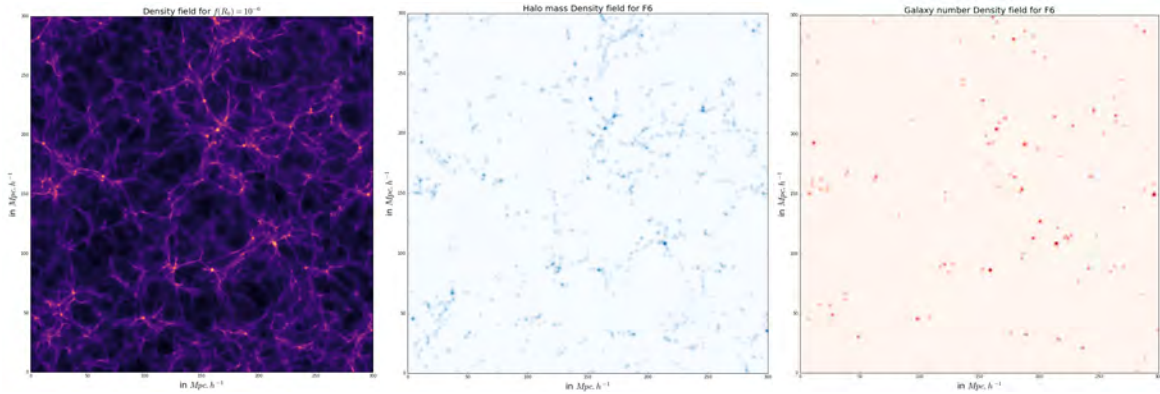


Figure 3: **Left panel:** Distribution of the Dark Matter density in a $10 \text{ Mpc}/h$ slice width from a N-body simulation. Brighter the color denser the dark matter is. **Center panel:** Distribution of the dark matter halos with mass greater than $10^{12} M_{\odot}/h$ in the same slice of the simulation. We can see that the halos live in the overdense region of dark matter distribution. In fact, more massive the halo denser the dark matter in average. This dependence refers to the halo bias. **Right panel:** Galaxy distribution (using Luminous Red Galaxies (LRGs)) from the same slice of the simulation. We can see that only part of the halos are populated by these galaxies and in fact correspond basically to the more massive ones when considering LRGs.

expansion and lead to the local collapse. When considering late Universe (after the Cosmic Microwave Background), the collapse stops when the particles velocities balance the gravitational potential. It is known as the Virial theorem which postulates that the initial velocities are randomly distributed. If it occurs at the very beginning of the Universe, when the particles do not have initial peculiar velocities, density increase enough in order to form a primordial black hole (instead of forming a virialized structure). However, due to the initial coupling between matter, photons and neutrinos break this condition and the late collapses will form dark matter halos embedding the normal matter and allowing the galaxy formation. It is crucial to notice that the history of galaxy formation is directly connected to the dark matter halos formation history. On an other hand, the astrophysics of the star formation and other processes responsible to the galaxy morphology are strongly non-linear and then difficult to link to the dark matter halo history; that is known as the "Galaxy - (Dark Matter) Halo connection". Be careful, that it is not exactly the name of the title of this note. The part of the Universe that will collapse and form the dark matter halos is not randomly distributed. As we can see on the figure 3, the massive halos form on the overdense part of the dark matter distribution. In other words, the distribution of the halos trace the most dense regions of the dark matter distribution. In fact, more massive the halo denser the average underlying dark matter field. This dependence refers to the halo bias that will be explain in more details in the section 2.

Finally the tracer bias is a combination of two components:

1. **Halo bias** : the bias of the halo hosting the galaxies we observe

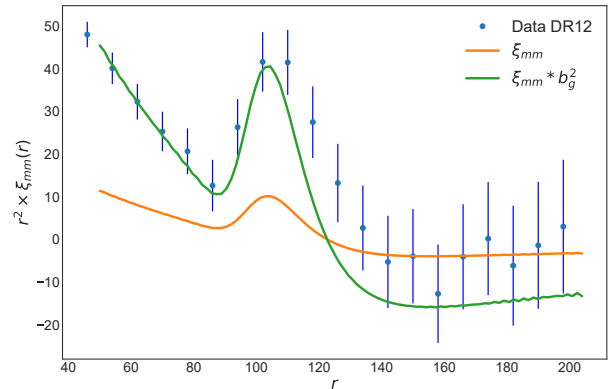


Figure 4: 2pt-correlation function from BOSS DR12 data Luminous Red Galaxies compare to the theoretical 2pt-correlation function of the matter $\xi_{mm}(r)$ with and without the galaxy bias. We can see that effectively the galaxy bias allows to obtain similar amplitude than the one obtain from galaxy data.

2. **Galaxy - Halo connection** : how the galaxies populate these halos (number of galaxies per halo for example).

That is fundamental in order to compare any model to the observable (i.e. tracers). We will now introduce in more details the concept of halo bias.

2 Dark Matter Halos

The dark matter halos formation is a corner stone at the interface between Astrophysics and Cosmology. One can often listen that Galaxy Clusters are larger virialized structure of the Universe. In fact what we call Galaxy Cluster are massive dark matter halos in which were formed various galaxies (luminous enough to be



Figure 5: **Left panel:** Spiral galaxy UGC-2885 (also known as Rubin's galaxy) image from the Hubble Space Telescope. **Right panel:** Elliptical galaxy M87 (also known as Virgo A) image from the Hubble Space Telescope.

observed). The galaxies formed because dark matter halos collapsed that allowed the formation of stars and so galaxies.

2.1 Field, group and cluster galaxies

You probably also listen about the distinction between field galaxies and group galaxies (living inside galaxy clusters). The difference is observational: an observed lonely galaxy is referred to a field galaxy, when few galaxies (5-10 or less) are observed together we refer to group galaxies and when we observe more galaxies in the structure we refer to cluster galaxies. The fundamental reason of this separation is related to the morphology of these galaxies. Indeed, the proportion of spiral galaxies is much higher in the field galaxy population when the elliptical galaxies dominate the cluster galaxies. This difference is due to the galaxy morphology evolution in part driven by galaxy collisions and mergers. A spiral galaxy can be seen as a young galaxy that constantly produces new stars by accretion process inside the disk (can be referred to **late-type galaxy** for this reason). For the star formation, the color of these galaxies tends to be blue and it is common to listen about **blue-type galaxies**. We present a nice image from the Hubble Space Telescope of the UGC-2885 spiral galaxy. When two spiral galaxies collide and eventually merge the disks are disrupted, the gas is efficiently consumed to form a burst of stars. The velocities tend to be randomly distributed that transform into an elliptical bulge. You can see a nice simulation of spiral galaxy merging clicking on this link : [Video link](#)². After the merging there is very little gas to transform into new stars, reason why the color seems to be redder than for a spiral galaxy. You probably inferred that the elliptical galaxies are often called **red-type galaxies**; and due to the fact that there is no more new star formation they appear to be populated by older stars and so were supposed to be formed at early time and are also called **early-type galaxies**. It exists a huge variety of morphology between the

²In case you can not open the hyperlink, the video direction is <https://youtu.be/CcopGM8iUCY>

two extremes presented that you can find in the Hubble diagram and we can also include Active Nuclei Galaxies and Quasars are other kind of galaxies that are useful for tracing the matter distribution.

Due to the collision history needed to form elliptical galaxies one can understand that it will be much more common to observe red-type galaxies inside massive groups and clusters. So we can understand that the elliptical galaxies will trace the densest regions of the Universe.

2.2 Halo Bias

The last sentence reflects an intuition and one can formalize and demonstrate it. First of all, we have to remember that few decades ago there were two proposed paradigms in order to explain the structure formation:

1. **Top-Down** : proposed that large scale overdensities collapse first and then fragment into smaller ones by instabilities process.
2. **Bottom-Up** : proposed (and is the correct one) that first the small scale structures collapse, forming galaxies in local small halos and then these small halos are absorbed in bigger halos that collapse later.

The presence of the Cold Dark Matter (CDM) explains why the second scenario is the correct one and allows us to understand the differences between the field and group galaxies as explained in section 2.1.

We present in figure 6 a 1D visualization of a Gaussian Random Field that mimics matter density field without Cold Dark Matter (left panel) and with Cold Dark Matter (right panel). The two horizontal lines represent the threshold of collapse for 2 different redshift values ($z=4$ and $z=2$). The filled regions represent the ones that have collapsed, in cyan for $z=4$ and blue for $z=2$. The mass of the collapsed structure is proportional to the width of the filled region. We can see that the presence of the CDM allows the fluctuations at smaller scale and that we first form smaller halos than in the case without CDM. The galaxies form

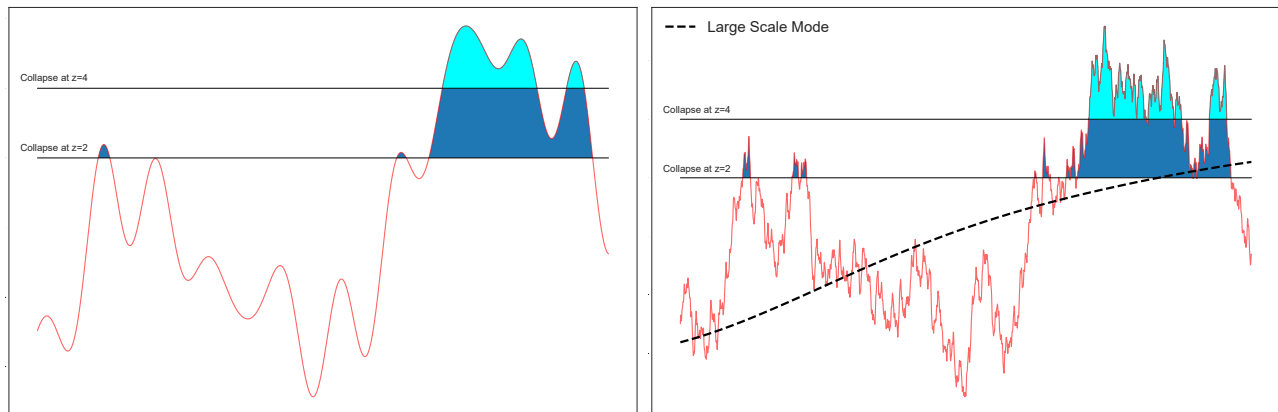


Figure 6: **Left panel:** Gaussian Random Field that mimic a Universe dominated by large scale perturbations like in absence of Cold Dark Matter. The two horizontal lines represents the threshold of collapse for 2 different redshift values ($z=4$ and $z=2$). The filled regions represents the one that have collapsed, in cyan for $z=4$ and blue for $z=2$. The mass of the collapsed structure is proportional to the width of the filled region. **Right panel:** The same Gaussian Random Field but including the small scales fluctuations that mimic the Universe we observe. We can see that we first form small structures at $z=4$ which will be regroup inside a larger halo at $z=2$. That's exactly the meaning of the Bottom-up paradigm. Moreover, we present the very large scale dominant mode in dashed line that represents the matter density at large scale. We can see that the most massive halos (i.e. filled area with larger width) are located on highest values of the large scale mode. That is the representation of the halo bias: the most massive halos are more biased because live in average in very dense large scale environment.

inside the small halos before to merge together inside larger halo : the small halos in cyan at $z=4$ merge together and are observed in the same large halo at $z=2$ (in blue). The halos without CDM are larger and the matter inside should fragment to form the individual galaxies. We know that the Bottom-up scenario is the correct one thanks to galaxy survey and precise detection of temperature fluctuations in the Cosmic Microwave Background (CMB). In the right panel we also present the value of the density field at large scale (the Fourier modes for very small k values). We can see that there is a correlation between the quantity of halos and how massive they can be and the value of the large scale modes. This is exactly what characterize the halo bias. More massive the clusters, higher the mean density at large scale (so the environment). This concept was formalized using Excursion Set method and then Peak Background Split (PBS) approach [3, 4].

2.3 Bias model

Indeed, we can predict from the CMB temperature map the quantity of galaxy cluster of given mass that will be formed at a given redshift that corresponds to the Halo Mass Function (HMF). The work is based on linear perturbation theory coupled to non-linear spherical and ellipsoidal collapse theory in the context of Gaussian Random Field. The derivatives of the HMF can be related to the halo bias parameters as presented in [3]. Another approach is to use higher order perturbation theory and develop the math expanding the bias terms coupled to the expansion of the density field and ob-

tain an expression for the biased power spectrum and 2pt-correlation function as shown in [2]. We will not develop any calculation here due to the complexity of the work mentioned and encourage the interested reader to directly read these nice papers.

3 Galaxy - Dark Matter Halos

We present the bias of the halos in the previous section, but as we already mentioned that we do not really observe the halos but galaxies. We need to understand which kind of galaxies populate which kind of halos at different time (i.e. different redshifts). It exists different kind of approach but we will focus on the two most often used:

1. **Halo Occupation Distribution (HOD)** : that propose to provide the mean number of a kind of galaxies and the mass of the hosting halo.
2. **Conditional Luminosity Function (CLF)** : that propose to estimate the density of galaxy (in a sort the number of galaxies) that posses a luminosity in function of the mass of the hosting halo.

3.1 Halo Occupation Distribution (HOD)

The HOD is the evaluation of the mean number of galaxies function of the halo mass. The idea is to compare the 2pt-correlation function measured from a sample of galaxies to the theoretical 2pt-correlation function of the halos convoluted with the mean number of

galaxies supposed to populate these halos. In fact, the HOD propose a functional that can mimic the global form of the mean number of galaxies function of the halo mass ($\langle N_{gal}(M) \rangle$). In fact, due to the bottom-up scenario, the galaxy integrate larger halos during the history by accretion process. For this reason, the galaxy that was formed at the center of the future larger halo will have a different way to evolve than the others. This central galaxy will experience higher rate of collision and sometimes merge with the other galaxy. For this reason this galaxy can grow much more than the other galaxies and have specific properties like the luminosity (see section 3.2). The mean number of galaxies is then expressed as the sum of the probability that the central galaxy is of the kind of galaxy, $\langle N_{cen}(M) \rangle$ you consider (so a value between 0 and 1) and the mean number of satellite galaxies (all minus the central) $\langle N_{sat}(M) \rangle$. A common functional form for the Luminous Red Galaxies is given by [8]:

$$\langle N_{tot}(M) \rangle = \langle N_{cen}(M) \rangle + \langle N_{sat}(M) \rangle, \quad (24)$$

$$\langle N_{cen}(M) \rangle = \frac{1}{2} \left[1 + \text{erf} \left(\frac{\log M - \log M_{min}}{\sigma_{\log M}} \right) \right], \quad (25)$$

$$\langle N_{sat}(M) \rangle = \left(\frac{M}{M_{sat}} \right)^\alpha \exp \left(-\frac{M_{cut}}{M} \right) N_{cen}(M), \quad (26)$$

that is fully describe by 5 free parameters (M_{min} , M_{cut} , M_{sat} , α , $\sigma_{\log M}$) to fit comparing with the target galaxy 2pt-correlation function. We present in figure 7 the evolution with the halo mass of the HOD fitted on the CMASS LRG sample (BOSS collaboration). We can see that the relation between the mean number of LRG galaxies with the halo mass is not trivial what lead to a non trivial relation between the density of LRG and the underlying large scale matter density field.

3.2 Conditional Luminosity Function

The HOD is a very useful information to link the number density of galaxies to the host halo mass. However, we can use other information like the luminosity of the galaxies. The luminosity is the intrinsic brightness of an object, that means we have to correct the dimming of the received flux by the distance and the effects of the Universe expansion. If the work is well done, the luminosity is comparable quantity between different galaxies at different redshifts and so different distances.

Using the lensing information from background galaxies it is possible to estimate the position and the mass of the group and clusters. Then, one can stack the information of the luminosity distribution of all the groups/clusters with similar mass and produce the distribution presented in figure 8. The results presented are inferred from the nice work done in [7]. It exists more recent works about CLF like [1, 6, 5]. The same work is also done for blue type galaxies and can also be used. For the same reason explained for the HOD

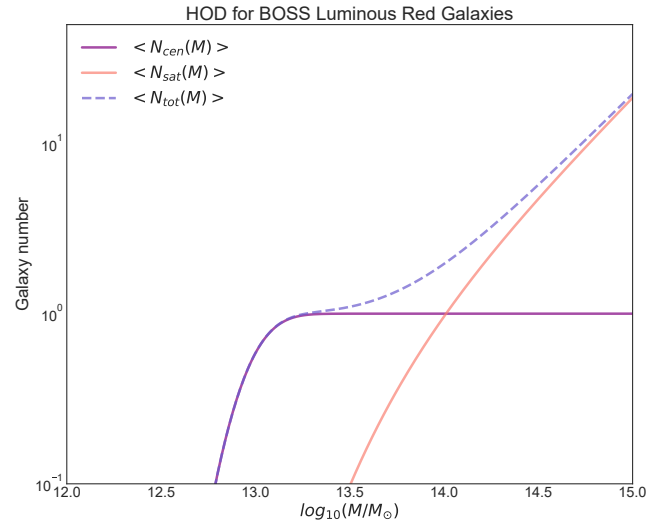


Figure 7: HOD for Luminous Red Galaxies (brightest elliptical galaxies) obtained using the BOSS DR12 sample for CMASS. This HOD use 5 free parameters to fit on the data. We can see that these LRGs only live inside halos of mass more massive than $10^{12.5} M_\odot$. We can also see that almost all the halos with a mass greater than $10^{13} M_\odot$ host a LRG as central galaxy. Finally, the number of LRG satellite galaxies increase with the halo mass. The most massive halos can host dozens of LRGs.

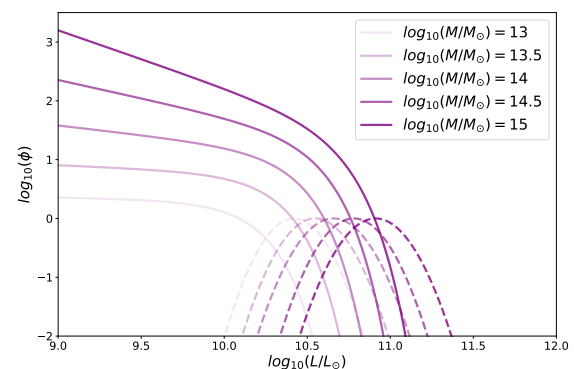


Figure 8: CLF for the red galaxies from [7] for different halo masses. The result is provided separately for the satellite galaxies and the central galaxy in order to take into account the cannibalism processes that occurs more often for the central ones.

construction one should separate the central galaxy and satellite galaxies for taking into account the difference of evolution processes. We expect the distribution of luminosity of the satellite galaxy to follow in some way the halo mass function for small halo masses, when the halo is supposed to form only one big galaxy (like the cyan halos in the figure 6). It should be the case if the stellar mass to halo mass ratio was a constant but the star formation process is complex and strongly depends on the environment. For this reason, the luminosity function for the satellite galaxies is expected to be a modified Schechter function (Schechter function describing the halo mass distribution form). On the other hand, the randomness of collision and merging processes let supposed that the luminosity of the central galaxy should tend to a log-normal one (Central Limit theorem). On an equivalent way than for the HOD, 5 free parameters are fitted on each mass range of the data in order to obtain the CLF. We present the CLF of red-type galaxies inferred for various halo masses from [7].

4 Conclusions

From both number density and luminosity information one can try to map the relation from galaxy observations and average most likely mass of the hosting halo. Combining the halo mass probability and the halo bias model, one can try to model the bias of the galaxy tracer. However, the complexity of the different connections it is important to keep in mind that the prediction can not be perfect and do not over interpret the final results from galaxy clustering if the effects measured are not so high. That was we usually call part of the systematics effects. We continuously try to reduce them and tracer bias start to be an important one for the further galaxy surveys.

References

- [1] Ting-Wen Lan, Brice Ménard, and Houjun Mo. “The galaxy luminosity function in groups and clusters: the faint-end upturn and the connection to the field luminosity function”. In: MNRAS 459.4 (July 2016), pp. 3998–4019. DOI: 10.1093/mnras/stw898. arXiv: 1509.07870 [astro-ph.GA].
- [2] Patrick McDonald and Arabindo Roy. “Clustering of dark matter tracers: generalizing bias for the coming era of precision LSS”. In: 2009.8, 020 (Aug. 2009), p. 020. DOI: 10.1088/1475-7516/2009/08/020. arXiv: 0902.0991 [astro-ph.CO].
- [3] Ravi K. Sheth and Giuseppe Tormen. “Large-scale bias and the peak background split”. In: MNRAS 308.1 (Sept. 1999), pp. 119–126. DOI: 10.1046/j.1365-8711.1999.02692.x. arXiv: astro-ph/9901122 [astro-ph].
- [4] Ravi K. Sheth and Giuseppe Tormen. “On the environmental dependence of halo formation”. In: MNRAS 350.4 (June 2004), pp. 1385–1390. DOI: 10.1111/j.1365-2966.2004.07733.x. arXiv: astro-ph/0402237 [astro-ph].
- [5] Chun-Hao To et al. “RedMaPPer: Evolution and Mass Dependence of the Conditional Luminosity Functions of Red Galaxies in Galaxy Clusters”. In: ApJ 897.1, 15 (July 2020), p. 15. DOI: 10.3847/1538-4357/ab9636. arXiv: 1910.01656 [astro-ph.GA].
- [6] M. Trevisan and G. A. Mamon. “A finer view of the conditional galaxy luminosity function and magnitude-gap statistics”. In: MNRAS 471.2 (Oct. 2017), pp. 2022–2038. DOI: 10.1093/mnras/stx1656. arXiv: 1705.10808 [astro-ph.GA].
- [7] Xiaohu Yang, H. J. Mo, and Frank C. van den Bosch. “Galaxy Groups in the SDSS DR4. II. Halo Occupation Statistics”. In: ApJ 676.1 (Mar. 2008), pp. 248–261. DOI: 10.1086/528954. arXiv: 0710.5096 [astro-ph].
- [8] Zhongxu Zhai et al. “The Clustering of Luminous Red Galaxies at $z \sim 0.7$ from EBOSS and BOSS Data”. In: ApJ 848.2, 76 (Oct. 2017), p. 76. DOI: 10.3847/1538-4357/aa8eee. arXiv: 1607.05383 [astro-ph.CO].

LOS IONES NEGATIVOS COMO PRUEBA CRÍTICA DE LA MECÁNICA CUÁNTICA

Guillermo Hinojosa Aguirre
Instituto de Ciencias Físicas de la
UNAM Campus Morelos

Cuándo la atracción de Coulomb no existe, explicar la unión de un electrón con un átomo es complejo. Otras fuerzas tienen que entrar en acción. Cuáles son estas fuerzas y cómo es posible que un ion negativo sea estable son preguntas cuyas respuestas han llevado a varios descubrimientos y aplicaciones. Una de las aplicaciones más conocidas es el método de datación con ^{14}C . Uno de los descubrimientos más recientes se hizo aquí en el Instituto de Ciencias Físicas de la UNAM.

Gracias a la espectrometría de masas es posible separar las componentes moleculares y atómicas de una muestra. Esta separación se hace como función del cociente entre su masa y su carga (m/q). Por ejemplo $m/q = 74.92$ indicaría la presencia de arsénico en una muestra de comida ($^{75}\text{As}^+$, por la carga $q = +1$). El mineral de uranio tiene al menos seis isótopos, uno de los cuáles es muy importante por su aplicación en la generación de energía verde. Gracias a la espectrometría de masas se sabe que el ^{235}U tiene una abundancia natural de 0.0072, mientras que el ^{238}U tiene una abundancia de 0.9927. En el caso del carbón, los isótopos que se usan para fechar las muestras históricas son el ^{12}C y el ^{14}C . El problema es que todas las muestras biológicas también contienen ^{14}N . Entonces $m/q = +14$ para el ^{14}C y para ^{14}N ¡no se podrían distinguir! El truco para separarlos es usar iones negativos. Resulta que el nitrógeno no forma iones negativos, en cambio el carbón sí. Con esta idea Willard Libby ganó el premio nobel de química en 1960 y estableció los campos de la paleontología y arqueología físicas.

Los iones negativos son extraños pero no raros. Son extraños porque tienen propiedades peculiares que los distinguen de los iones positivos y de los átomos. Una de esas propiedades es que la energía de ligadura del electrón extra es baja, por ejemplo del orden de ~ 3 eV para los alógenos. A la energía de ligadura del electrón extra se le conoce como afinidad electrónica (AE). Otra propiedad es que el potencial de enlace del electrón es de muy corto alcance (r^{-5}). La interacción entre los electrones cobra especial importancia en la estabilidad de estos iones. No todos los átomos forman o tienen iones negativos estables. Por ejemplo los gases nobles y el nitrógeno, de lo que Libby se dió cuenta.

No son raros porque son sorprendentemente abundantes en la Naturaleza. Por ejemplo, en la atmósfera terrestre, en los plasmas y hasta en el espacio interestelar. En este último ambiente son especialmente difíciles de justificar por su fragilidad pero sobre todo porque no se conoce su mecanismo de formación.

Como el electrón extra está débilmente ligado [10] al átomo es razonable esperar que éste lo pierda fácilmente, por ejemplo, por

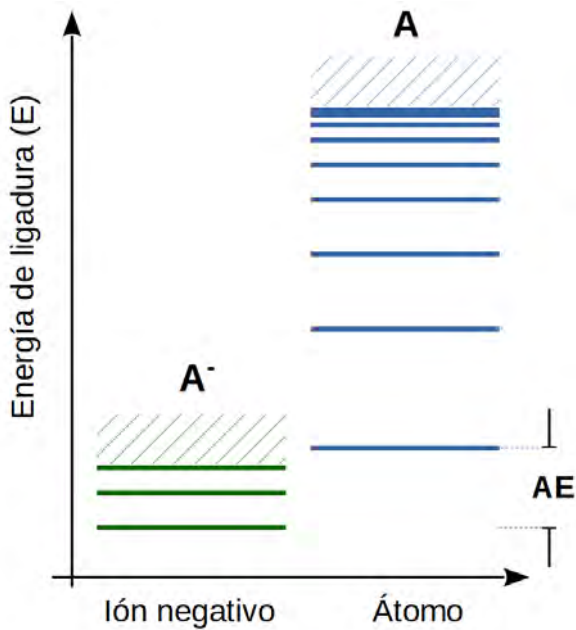


Figura 1: Niveles de energía de un átomo (A) comparados contra los niveles de energía del anión correspondiente (A^-) (no a escala). Las líneas horizontales corresponden a la energía de los estados base (las más bajas) y a algunos estados excitados. Los niveles de energía se apilan cada vez más cerca conforme aumenta la energía (como $1/n^2$ donde n es el número cuántico principal) hasta que un electrón se desliga del átomo y se pierde en el continuo (región de líneas oblicuas). La afinidad electrónica AE corresponde a la ecuación (4).

interacción con otros átomos o con fotones. Cerca de las estrellas hay radiación y alta probabilidad de colisiones, sin embargo, la misión espacial Cassini [1] encontró que los iones negativos son más abundantes de lo esperado. Además estas cantidades de aniones o iones negativos no se pueden justificar con la teoría existente [2]. De hecho, no existe hasta ahora una teoría o un modelo que justifique la cantidad de iones negativos en el espacio interestelar.

Otro aspecto interesante de los iones negativos es la aparente contradicción entre el hecho de que la correlación entre electro-

nes es necesaria para justificar su estabilidad, pero al mismo tiempo, en colisiones el electrón extra se puede describir como un electrón libre. Ambos aspectos son fundamentales para describir a los iones negativos.

El caso del H^- , se trata de una referencia para el estudio de iones negativos. Es isoelectrónico con el He pero son esencialmente diferentes. El H^- es una estructura que es posible gracias a una combinación entre blindaje (o el apantallamiento del campo eléctrico) electrostático y polarización. A principios del siglo XX el origen de la mayoría de la absorción en el espectro visible de las atmósferas de estrellas frías era uno de los principales problemas en astrofísica. Los astrónomos creían que el origen del continuo en la opacidad en la zona visible del espectro se debía a la presencia de átomos de metales. Sin embargo, deberían tener abundancias irreales. Fue entonces cuando Wild en 1939 [3] calculó el coeficiente de absorción del H^- y demostró que este ión domina la opacidad en el espectro visible en estrellas como el Sol a pesar de que su ionósfera puede estar a varios miles de grados de temperatura.

FUNCIONES DE ONDA PARA EL H^-

Surge entonces la curiosidad sobre la estructura y sobre los procesos físicos que pueden formar esta clase de iones. Primero, estudiemos cómo puede ser una función de onda para el H^- . Para ello me basaré en la tesis doctoral de la estudiante de mi amigo Marcelo de Brazil [4] que tiene una introducción muy buena sobre este tema. A continuación reproduzco una parte de su tesis.

Las funciones de onda ψ representan la posición de un electrón en el espacio (en este caso). Estas funciones de onda se pueden usar para derivar *observables* que pueden ser la posición promedio o la energía. En

la práctica, se aplican métodos aproximados para encontrar estos valores esperados u observables. En el caso del H^- las funciones de onda deben poder describir a dos electrones. Como un inicio, tomemos una función de onda y *minimicemos* con algún método aproximado la energía nuestro H^- (que se conoce como *hidruro* entre los químicos). Durante este ejercicio suponemos que siempre usamos el mismo método de solución. Sea,

$$\psi_0(r_1, r_2) = e^{(Z - \frac{5}{16})(r_1 + r_2)} \quad (1)$$

en donde r_1 y r_2 son las posiciones de los electrones 1 y 2 respectivamente. Z es el número atómico que en este caso es $Z = 1$ para el hidrógeno. Nótese que a diferencia del átomo de helio, que también tiene dos electrones, los electrones del H^- no son indistinguibles. Se trata de dos electrones diferentes. De esta ψ la energía que se obtiene es -0.473 Ry (o -12.866 eV). Hay que tomar en cuenta que la energía de ligadura o la energía del estado base del H^- es menor que la energía del átomo de hidrógeno (-13.601 eV). Como la energía que se obtiene para el ión es mayor que la del átomo, entonces esta función de onda no logra predecir la estabilidad del anión. La energía de los aniones debe ser menor que la del átomo correspondiente (véase la Fig. 1). Además, esta función de onda no toma en cuenta el hecho de que los electrones son distinguibles. Al contrario que en el caso del He.

Una función de onda que toma en cuenta el carácter distinguible de los electrones del H^- es:

$$\psi_1(r_1, r_2) = N(e^{-r_1 - \beta r_2} + e^{r_2 - \beta r_1}) \quad (2)$$

en donde β es un parámetro que toma en cuenta el blindaje del electrón 1 por el electrón 2. En forma más general,

$$\psi_2(r_1, r_2) = N(e^{-\alpha r_1 - \beta r_2} + e^{-\alpha r_2 - \beta r_1}) \quad (3)$$

los parámetros α y β se pueden calcular al minimizar la energía y dan $\alpha = 1.039$ y $\beta = 0.283$. Con estos valores, la energía del estado base del H^- da -0.513 Ry que sí predice la estabilidad del H^- pero aún no predice un valor exacto para la energía. Esta última función de onda, en comparación con las anteriores, considera la correlación entre los electrones. En conclusión, no es posible predecir la existencia de este ión negativo sin considerar el amarramiento de los dos electrones.

Un parámetro que se usa para describir a los iones negativos en la *afinidad electrónica* (AE) y que se define como la diferencia de energía entre la energía del estado base del átomo E_A y la energía del estado base del anión correspondiente E_{A^-} ,

$$AE = E_A - E_{A^-} \quad (4)$$

Hasta ahora, las funciones de onda como la ecuación (3) logran predecir la existencia del ión negativo, pero fallan en predecir el valor de la AE medida, que en el caso del H^- es 0.754 eV. La función de onda propuesta por Chandrasekhar es,

$$\psi_3(\vec{r}_1, \vec{r}_2) = \psi_2(r_1, r_2)(1 + c|\vec{r}_1 - \vec{r}_2|) \quad (5)$$

que al minimizar la energía, arroja $\alpha = 1.075$, $\beta = 0.478$ y $c = 0.312$. Para la energía, resulta $E = -0.526$ Ry (o -13.305 eV). Si aplicamos la eq. (4) este resultado arroja que $AE = 0.704$ eV. Que aún difiere del valor medido para AE pero no está lejos. De los valores para las constantes α y β obtenidos de minimizar las energías para las funciones de onda de las ecuaciones

(3) y (5) se concluye que el apantallamiento del campo eléctrico del primer electron, queda casi sin cambios. En cambio el efecto del término $(1 + c|\vec{r}_1 - \vec{r}_2|)$ consiste en disminuir el apantallamiento o blindaje en el segundo electrón. Esto indica que el H^- está fuertemente polarizado-

EL CONCEPTO DE TIEMPO DE VIDA

Hasta aquí se entiende que un anión es estable cuando la AE es positiva. Es decir, que la energía del estado base del ión negativo es menor que la energía del estado base del átomo (o molécula) correspondiente; esto hace aún más sorprendentes a los iones negativos. La energía es una buena forma de saber si un anión es estable pero además existe un parámetro interesante que se conoce como la *vida media*, $T_{1/2}$ o el *tiempo de vida* que en física se conoce como τ . Es una medida de la velocidad con la que una cierta cantidad partículas decae o cambia durante el tiempo y se define como,

$$A_f = A_0 e^{-t/\tau} \quad (6)$$

en donde t es el tiempo, A_f es la intensidad o población final e A_i es la cantidad inicial de la población. Entonces, τ es el tiempo en que una población de una especie decae de una cierta cantidad arbitraria, digamos A_i , a A_i/e . Como es $e \approx 2.72$, se puede decir que el tiempo de vida es el tiempo en el que una cantidad determinada de una población disminuye a un tercio, más o menos.

Aquí es un buen momento para pensar lo que dos investigadores dijeron en una conferencia (yo estaba ahí):

- *el tiempo que ha transcurrido es mayor que el tiempo de vida, por lo tanto, la cantidad de partículas debe ser nula, ya no hay*

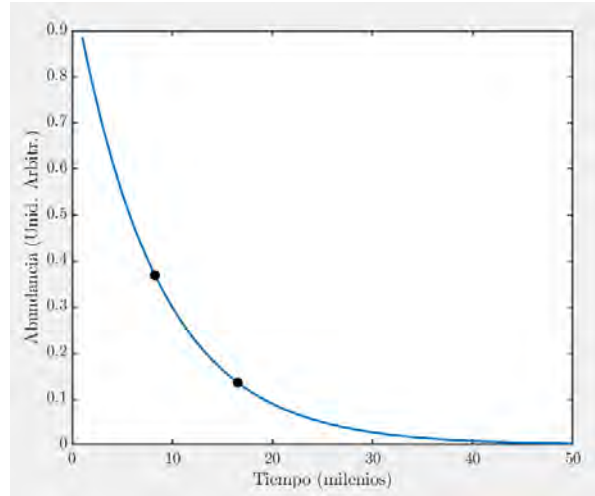


Figura 2: El concepto de tiempo de vida τ de la ecuación (6). Los puntos negros están separados por un tiempo $t = \tau$. En el caso del ^{14}C , la cantidad del isótopo de masa 14 se mantiene constante. Cuando una planta o un ser vivo dejan de intercambiar carbono con el medio ambiente, entonces la cantidad de ^{14}C decae sin *renovarse* con el intercambio con el medio ambiente. Al cabo de 8266.61 ± 58 años más o menos la tercera parte de este isótopo de desintegra comparado con su abundancia natural. Hay correcciones, por ejemplo, hay una corrección de origen antropomórfico debido a las pruebas de armas nucleares entre 1940 y 1990. Está también la explicación de el por qué la proporción $^{14}C:^{12}C$ se mantiene constante en el ambiente. Es un tema muy interesante.

Deténgase un momento para pensar porqué esta expresión es conceptualmente incorrecta.

Los iones negativos también decaen con el tiempo de manera exponencial según la ecuación (6). Por ejemplo, el anión molecular que se forma cuando el conocido *vinylene* captura un electrón, CCH_2^- , tiene un tiempo de vida medido de 102 s. El estado excitado $(2p^2)^3P$ del H^- parece que tiene un tiempo de vida del 44 ± 4 ns. Mientras que el tiempo de vida del estado base del H^- el $(2s^2)^1S$ no se sabe cuanto es. Está bien establecido el hecho de que su τ debe estar -al

menos- en la escala de los μs pero nunca se ha podido medir y creo que ni calcular.

Hay especies atómicas y moleculares que no forman iones negativos estables. Como ya se mencionó el nitrógeno no genera un anión estable, al igual que los gases nobles. Por ejemplo el neón. El neón tiene diez electrones y se dice que es isoelectrónico con el metano porque ambos tienen diez electrones y sus orbitales son análogos. El Ne por ser un gas noble no forma iones negativos y se creía que el metano siendo isoelectrónico con el neón, tampoco formaría iones negativos.

El metano CH_4 es la principal componente del gas natural. El gas natural es la principal alternativa al carbón mineral ya que no produce CO_2 . La sustitución del carbón mineral por gas natural es el primer frente de batalla contra el calentamiento global en las termoeléctricas de ciclo combinado. Sin embargo, el metano produce un efecto invernadero sesenta veces más potente que el dióxido de carbono si se libera en la atmósfera sin deflagarse.

La pregunta de si el metano forma iones negativos es importante porque los modelos de plasmas atmosféricos en donde hay metano o en los modelos de flamas (las flamas de metano son muy comunes en las refinerías) se tienen que corregir. No vamos a ahondar en la importancia de anión del metano, solo enfatizo que su demostración sería uno de los avances más importantes en el campo, en las últimas décadas.

LOS INDICIOS

Como se mencionó, los espectrómetros de masas distinguen iones como función de su relación entre su masa y su carga (m/q). El experimento que usamos en el ICF tiene como primera fase un espectrómetro de masas. Ahí observé una señal altísima en

$m/q = -16$ sin inyectar oxígeno. El ión de O^- también tiene una relación de masa carga de -16. Después se presentaron una serie de situaciones que no se podían justificar con la presencia de alguna contaminación de oxígeno:

- La señal estaba correlacionada con la inyección de metano.
- El sistema se puede mantener muy limpio de contaminación ya que la cámara del plasma está fabricada con cuarzo.
- La señal no disminuyó como función del tiempo, que es lo que se espera cuando se trata de una contaminación con oxígeno.
- Los gases que se usaron son de ultra alta pureza. La señal en $m/q = -16$ no tenía la intensidad de una traza.

Lo que nos llevó a hacer una investigación rigurosa del asunto [5]. En esa investigación encontramos aún más evidencia, que sumo sin mucho detalle:

- La intensidad de la señal en $m/q = -16$ dependía de la presión en la cámara de plasma. Esto indica que se trata de una especie propia del plasma y no un contaminante.
- La sección transversal de despojo electrónico no coincidía con la sección para el O^- , o sea que no se trataba de oxígeno.
- El radio promedio del ión arrojó un valor extraordinariamente alto, de 3.5 \AA que no tiene nada que ver con el valor correspondiente al O^- que es de $\approx 1 \text{ \AA}$.

Además había cierta evidencia. Primero ¿cómo era posible que nunca reportaran masa 16 en los plasmas de metano? Vaya, ni como contaminante [6]. Pero sí hay quienes están dispuestos a pagar el precio de la verdad. Por ejemplo, Winter y sus colaboradores [7] reportaron en un estudio de plasmas con metano: *...five minutes later we added a constant flow rate of 2 sccm CH₄ to the reaction chamber resulting in an increase in the mass spectrometer signal at mass 16 and mass 26...* Evidentemente lo vieron, pero no lo asignaron. Incluso más adelante reportan: *As these particles can survive in the discharge for more than 5 min (up to 20 min in one experiment) they have to be charged negatively ...* Eso fue música para nuestros oídos.

Otra pieza de evidencia es de unos autores que no se atrevieron ni a mencionarlo. Pero se vieron forzados a reportarlo por lo menos es sus gráficas de datos porque eso es lo que midieron. Fue Hollenstein y sus colaboradores [8]. Me atrevo a reproducir aquí una parte de su gráfica (ver Fig. 3). No lo mencionan como tal, pero si pusieron sus datos abiertamente. Corresponde al picote de señal en el agregado de orden uno.

Aún más, encontramos otra pieza de evidencia. Boltz y Glick [9] esta vez con espectrometría de electrones (no de masas) observaron resonancias temporales que asignaron a metano negativo. Ellos lo atribuyeron a unas estructuras temporales que se forman durante un tiempo muy corto¹, del orden de 10^{-14} s. Debido al método que usaron, no se puede saber más sobre el tiempo de vida, pero sí reportaron su existencia.

Con toda esa evidencia en la mano, publicamos nuestros resultados. Los editores de la revista invitaron a tres revisores muy severos. Finalmente, se publicó y logramos

¹estados de Jahn-Teller

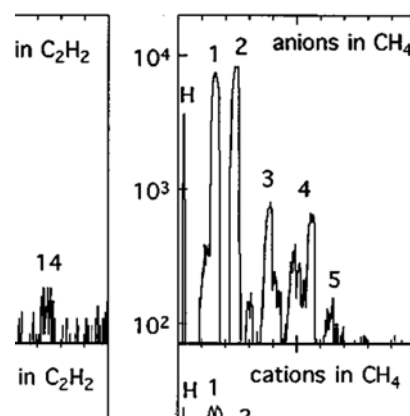


Figura 3: Una parte de la gráfica del espectrómetro de Hollenstein *et al.* [8]. Claramente observaron metano negativo. Para ellos se trataba del agregado de metano de orden 1.

que nuestros críticos desenvainaran sus espadas con la idea de cortarnos la cabeza. Los argumentos en contra de la existencia del metano son válidos. Aunque francamente no creo que supieran por qué. El argumento en contra de la existencia del metano se basa en el hecho de que el metano, con diez electrones, es isoelectrónico con el neón. El neón, siendo un gas noble de capa cerrada, no forma iones negativos estables y por lo tanto el metano, que también es de capa cerrada, tampoco. Pero nada puede contra la rigurosidad de una observación experimental bien hecha. El tiempo transcurrió y un día, el Dr. Ramirez-Solís llegó al laboratorio con una idea...

EL DESCUBRIMIENTO

¿Qué tal si el CH₄⁻ es isoelectrónico con el sodio y no con el neón? Finalmente, el metano negativo tiene once electrones, no diez. El sodio es bien conocido por tener unos estados muy excitados (las líneas azules en la Fig. 4) que tienen unos tiempos de vida ridículamente grandes.

Esa clase de estados existen gracias a que las reglas de transición atómicas prohíben que regresen, a través de la emisión de un fotón, al estado base. Como la transición contradice las reglas de conservación del momento angular, esos estados tienden a permanecer un cierto tiempo pero decaen *radiativamente* y siguen la ecuación (6). Se llaman estados metaestables y son, entre otras cosas, los responsables del funcionamiento de los láseres.

El Dr. Ramírez-Solís utilizó sus herramientas de análisis con la estructura molecular del metano negativo con la idea de ver si pasaba lo mismo que con el sodio, pero para el caso de la molécula del metano con once electrones. Fue increíblemente sorprendente, ver que la mecánica cuántica puede predecir, (en manos de alguien como el Dr. Ramírez-Solís) de primeros principios, la estabilidad del metano negativo y hasta ofrecer una estructura compatible. Los cálculos son complejos porque se trata de la evolución temporal de algo desconocido.

La estructura que resultó es conocida como un *exciplex* que en este caso es lineal (ver Fig. 5). El trabajo se publicó [11] y entonces lo que sigue es anunciar el descubrimiento en congresos, pero llegó la pandemia...

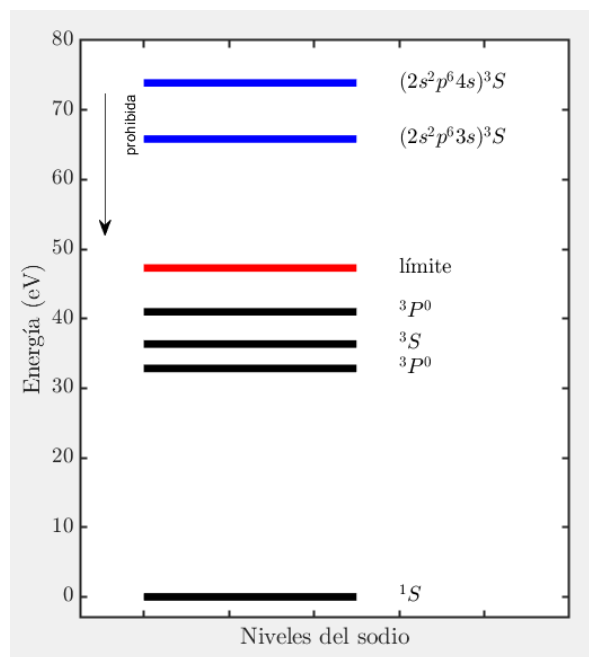


Figura 4: Algunos estados del sodio. Se han graficado solamente algunos pocos para fines de ilustración. La línea de color rojo indica el límite en donde un electrón pierde el enlace con el átomo y entonces el átomo se ioniza a Na^+ . Sin embargo, existen estados muy excitados (las líneas azules) en los que los átomos y las moléculas permanecen estables. Estos estados tienen la siguiente característica: su transición de regreso al estado base está prohibida por las reglas de transición de la mecánica cuántica.

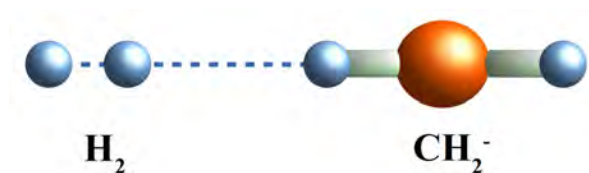


Figura 5: A temperatura ambiente y después de 8 ps, los cálculos muestran que se forma un estructura que se conocen como exciplexes. Esta molécula lineal es metaestable cuando $\text{CH}_2^-(s = 3/2):\text{H}_2(s = 0)$

Bibliografía

- [1] Thomas J. Millar, Catherine Walsh, and Thomas A. Field. Negative ions in space. *Chem. Rev.*, 117:1765, 2017.
- [2] Marjan Khamesian, Nicolas Douguet, Samantha Fonseca dos Santos, Olivier Dulieu, Maurice Raoult, Will J. Brigg, , and Viatcheslav Kokoouline. Formation of CN^- , C_3N^- , and C_5N^- molecules by radiative electron attachment and their destruction by photodetachment. *Phys. Rev. Lett.*, 117:123001, 2016.
- [3] R. Wildt. Negative ions of hydrogen and the opacity of stellar atmospheres. *Astrophys. J.*, 90:611–620, 1939.
- [4] Bárbara da Fonseca Magnani. *Íons negativos em colisões com átomos e moléculas a velocidades intermediárias e baixas*. PhD thesis, Universidade Federal do Rio de Janeiro Instituto De Física, Universidad de Federal de Rio de Janeiro, Brazil, 12 2014. Tesis doctoral dirigida por Dr. Marcelo Martins Sant’Anna.
- [5] E M Hernández, L Hernández, C Martínez-Flores, N Trujillo, M Salazar, A Chavez, and G Hinojosa. Electron detachment cross sections of CH_4^- colliding with O_2 and N_2 below 10 keV energies. *Plasma Science Sources and Technology*, 23:015018, 2014.
- [6] G. Drabner, A. Poppe, and H. Budzikiewicz. The composition fo the CH_4 plasma. *Int. Journal of Mass Spectrometry and Ions Processes*, pages 1–33, 1990.
- [7] J Winter, J Berndt, S-H Hong, E Kovačević, I Stefanović, and O Stepanović. Dust formation in Ar/CH_4 and $\text{Ar}/\text{C}_2\text{H}_2$. *Plasma Science Sources and Technology*, page 034010, 2009.
- [8] Ch. Hollenstein, W. Schwarzenbach, A. A. Howling, C. Courteille, J.-L. Dorier, and L. Sansonnens. Anionic clusters in dusty hydrocarbon and silane plasmas. *Journal of vacuum science and technology*, page 535, 1996.
- [9] F. K. Botz and R. E. Glick. Methane temporary negative ion resonances. *Chemical Physics Letters*, page 279, 1975.
- [10] Kiattichart Chartkunchand. *Photodetachment Studies Of Atomic Negative Ions Through Velocity-Map Imaging Spectroscopy*. PhD thesis, University of Nevada, Reno, 5 2015. Tesis doctoral dirigida por Dr. Aaron M. Covington.
- [11] Alejandro Ramírez-Solís, Jacques Vigué, Guillermo Hinojosa, and Humberto Saint-Martin. Solving the CH_4^- riddle: The fundamental role of spin to explain metastable anionic methane. *Phys. Rev. Lett.*, 124:056001, 2020.



Sobre las propiedades mecánicas de la luz

R. Jáuregui

Departamento de Física Cuántica y Fotónica
Instituto de Física, Universidad Nacional Autónoma de México
Cd. de México C.P. 04510, México
rocio@fisica.unam.mx

1 Introducción.

Nuestra intuición física primaria se adquiere desde la niñez con base en la observación cotidiana de la evolución en el tiempo de la posición y orientación de objetos que interactúan entre sí ya sea gravitacional y/o electromagnéticamente. Estos objetos se caracterizan por poseer directamente ó a través de sus componentes propiedades que se identifican como masa inercial, carga eléctrica y momento dipolar magnético. Estas observaciones pueden plantearse como efectos de una *acción a distancia* entre los cuerpos; la interacción muestra dependencia en la distancia relativa y, en ocasiones, también en la velocidad relativa entre los cuerpos. En el caso electrodinámico, la visión alternativa de James Clerk Maxwell reemplazó la descripción en términos de acción a distancia por una interacción *local* entre los objetos materiales y campos eléctricos \vec{E} y magnéticos \vec{B} . En el modelo electrodinámico de Maxwell, en regiones ideales cuasi-libres de campos electromagnéticos la materia evoluciona libremente y en regiones libres de fuentes de campo, el campo electromagnético también evoluciona libremente; en ambos casos las variantes en la configuración de campos o entes materiales se derivan de las condiciones a la frontera de las regiones espacio-temporales en que se localizan. Cuando ambos, campo electromagnético y materia están presentes la interacción queda codificada en forma local en el espacio y en el tiempo a través de un término de acoplamiento fundamental que, en la notación contemporánea, se escribe simplemente como¹

$$\sum_{\mu=0}^3 J_{\mu}(x)A^{\mu}(x) \equiv J_{\mu}(x)A^{\mu}(x), \quad x : (ct, \vec{x}), \quad (1)$$

con $J_{\mu} : (\rho c, \vec{j})$ la densidad de corriente asociada a la materia y $A^{\mu}(x) : (\phi(x), \vec{A}(x))$ el potencial electromagnético,

$$\vec{E} = -\frac{\partial \vec{A}}{\partial ct} - \vec{\nabla}\phi, \quad \vec{B} = \vec{\nabla} \times \vec{A}. \quad (2)$$

Esto es, los entes involucrados son la materia que genera y responde *mecánicamente* a los campos $(\vec{E}; \vec{B})$ y los campos $(\vec{E}; \vec{B})$ así mismo afectan *mecánicamente* a la materia que les da origen. Esta expresión para la interacción es válida también en el ámbito de una descripción cuántica de la materia mediante el uso de la expresión adecuada para la densidad de corriente. Tampoco pierde su validez cuando reconocemos al campo electromagnético como poseedor de propiedades que se especifican en términos de su estado cuántico.

La teoría electrodinámica es en principio conceptualmente sencilla, sus consecuencias son, sin embargo, espectaculares. Con ella, se mostró a los campos eléctricos y magnéticos como parte de un todo: el campo electromagnético (E.M.)

$$F_{\mu\nu} = \partial_{\mu}A_{\nu} - \partial_{\nu}A_{\mu} = \begin{pmatrix} 0 & -E_1 & -E_2 & -E_3 \\ E_1 & 0 & -B_3 & B_2 \\ E_2 & B_3 & 0 & -B_1 \\ E_3 & -B_2 & -B_1 & 0 \end{pmatrix}, \quad (3)$$

el cual satisface las ecuaciones de Maxwell,

$$\partial_{\rho}F_{\mu\nu} + \partial_{\nu}F_{\rho\mu} + \partial_{\mu}F_{\nu\rho} = 0, \quad (4)$$

$$\partial_{\mu}F^{\mu\nu} = \frac{4\pi}{c}J^{\nu} \implies \partial^{\mu}\partial_{\mu}A^{\nu} = \frac{4\pi}{c}J^{\nu} + \partial^{\nu}\partial_{\mu}A^{\mu} \quad (5)$$

y la luz, de velocidad c , se convierte en una manifestación de configuraciones ondulatorias de los campos E.M.². Es decir, aún en el contexto clásico, la teoría unificó la descripción de fenómenos que se pensaban como independientes. Adicionalmente, sin estar Maxwell conciente de ello, esta teoría da posteriormente pie a la concepción relativista de Einstein.

Maxwell y otros investigadores destacados de finales del siglo XIX y principios del siglo XX, asignaron al campo electromagnético propiedades mecánicas medibles en términos de sus efectos mecánicos sobre la materia. La electrodinámica describe a la interacción luz-materia en términos de intercambio de energía, momento lineal y momento angular. Este intercambio satisface reglas de conservación y sabemos, a través del teorema de Noether, que estas reglas de conservación pueden verse como un reflejo de la simetría ante traslaciones y rotaciones del espacio-tiempo subyacente.

¹Por lo general utilizaremos la notación según la cual se efectúan sumas cuando los índices en una expresión se repiten. Si estos índices son griegos corren sobre los valores 0,1,2 y 3 referidos a coordenadas espacio-temporales, y si son latinos de 1, 2, 3 refiriéndose a coordenadas espaciales cartesianas.

²A partir de la Ec. (5) podemos calcular el resultado de aplicar el operador $\partial^{\sigma}\partial_{\sigma}$ a $F_{\mu\nu}$ encontrando una ecuación de onda que se propaga a velocidad c y tiene como fuente a variaciones de la corriente J_{μ} . Adicionalmente en la norma de Lorentz $\partial_{\sigma}A^{\sigma} = \partial_{ct}\phi + \vec{\nabla} \cdot \vec{A} = 0$ y A_{ν} satisface la ecuación de ondas que tienen como fuente a $4\pi J^{\mu}/c$

El objetivo de este ensayo es doble. Por un lado comentamos brevemente sobre lo que ocurre en regiones del espacio cuando fronteras físicas rompen estas simetrías pero preservan otras. ¿Existen entonces cantidades mecánicas que permitan describir esta interacción en forma simple? ¿Se observan o no, reglas de conservación? No se pretende dar una respuesta de validez general sino mostrar mediante ejemplos específicos simetrías en las que identificamos a estas variables y comprobamos experimental y/o teóricamente su relevancia.

El segundo objetivo corresponde a hacer evidente la similitud que existe entre la formulación cuántica de las propiedades mecánicas – clásicas y cuánticas– de los objetos materiales descritas en términos de los operadores $\hbar(i\partial_{ct}, -i\vec{\nabla})$ y las propiedades mecánicas de los campos electromagnéticos basada en estos operadores. Esto es, mostraremos que, en condiciones especialmente importantes para la óptica cuántica, se puede definir un producto escalar para el campo A_μ a partir del cual se pueden evaluar las cantidades mecánicamente relevantes asociadas a luz estructurada con simetrías diversas, esto en total analogía con la dinámica cuántica de Heisenberg.

2 El tensor de energía momento y la identificación de un producto escalar para el campo de ondas electromagnéticas.

Con independencia del carácter clásico o cuántico que posea una densidad de corriente externa J_μ se puede mostrar que el tensor

$$\Theta^{\mu\nu} = \frac{1}{4\pi} \left[g^{\mu\nu} F^{\sigma\lambda} F_{\sigma\lambda} + F^{\mu\sigma} F_\sigma^\nu + g^{\mu\nu} J_\sigma A^\sigma - J^\mu A^\nu \right] \quad (6)$$

satisface la ecuación

$$\partial_\mu \Theta^{\mu\nu} = \frac{1}{4\pi} A_\mu \partial^\nu J^\mu \implies \partial_\mu T^{\mu\nu} = \frac{1}{4\pi} J_\sigma F^{\sigma\nu} \quad (7)$$

$$T^{\mu\nu} = \frac{1}{4\pi} \left[g^{\mu\nu} F^{\sigma\lambda} F_{\sigma\lambda} + F^{\mu\sigma} F_\sigma^\nu \right] \quad (8)$$

de tal suerte que en regiones del espacio tiempo en que no hay fuentes de campo electromagnético $T^{\mu\nu}$ es simétrico, invariante ante transformaciones de norma, tiene traza cero y da origen a cuatro-cantidades conservadas, $T^{0\nu}$, $\nu = 0, 1, 2, 3$. Explícitamente,

$$T^{00} = \frac{1}{8\pi} (\vec{E} \cdot \vec{E} + \vec{B} \cdot \vec{B}), \quad T^{0i} = \frac{1}{4\pi} (\vec{E} \times \vec{B})^i \quad (9)$$

corresponden a la densidad de energía y momento del campo electromagnético,

La variación de la energía del campo E.M. dada por la integral sobre cualquier volumen \mathcal{V} contenido en esas regiones solo varía en términos del flujo asociado al vector de Poynting y la energía mecánica transmitida a la corriente \vec{j} ,

$$\frac{d}{dct} \int_{\mathcal{V}} d^3x T^{00} = \frac{1}{c} \int_{S_{\mathcal{V}}} d\vec{S} \cdot (\vec{E} \times \vec{B}) + \int_{\mathcal{V}} d^3x \vec{E} \cdot \vec{j} \quad (10)$$

El vector de Poynting es proporcional a la densidad de momento del campo; este hecho fundamenta—cuando se cuantiza al campo E.M— a la relación $E = pc$ para un fotón.

Un cálculo directo muestra que la densidad de energía y a la densidad de momento quedan dadas en términos del potencial electromagnético A^μ por las expresiones,

$$T^{00} = \frac{1}{8\pi} \left[A_s \partial_{ct} E_s - E_s \partial_{ct} A_s \right] + \partial_s [\phi E_s + A_k [\partial_s A_k - \partial_k A_s]] + 4\pi [\phi \rho + \vec{A} \cdot \vec{j}] \quad (11)$$

$$T^{0i} = \frac{1}{4\pi} \left[E_s (-\partial_i A_s) + \partial_s [E_s A^i] \right] + 4\pi [\rho A^i] \quad (12)$$

siempre que se cumpla la norma de Lorentz

$$\partial_\mu A^\mu = \frac{\partial \phi}{\partial ct} + \vec{\nabla} \cdot \vec{A} = 0.$$

Las Ecs. (11-12) muestran que los generadores de traslaciones temporales, ∂_{ct} , y espaciales $\vec{\nabla}$ se relacionan directamente con estas cantidades conservadas. Estas expresiones, sin embargo, son dependientes de norma; el flujo de energía no está únicamente determinado por el vector de Poynting sino hay que incluir los términos que aparecen en la divergencia de la Ec. (11); un comportamiento análogo lleva a que el flujo de momento asociado a la Ec. (12) involucre a un tensor de esfuerzos dependiente de norma; finalmente, el intercambio con la materia de estas cantidades conservadas modificadas son también dependientes de norma.

Todo esto hace ver esta descripción para la conservación de energía y momento, como más complicada y alejada a nuestra intuición física. Sin embargo, da entrada a interpretaciones más cercanas a la formulación ondulatoria de la materia y, como veremos, puede simplificar cálculos relacionados con la propiedades de polarización de luz estructurada con simetrías identificadas.

Para hacer evidente el primer punto consideremos a una partícula escalar descrita por una función de onda compleja Ψ . Si esta función es estacionaria, $\Psi(x) = e^{-i\omega t} \psi(\vec{x})$, la condición de normalización usual se escribe como

$$\int d^3x \psi^* \psi = 1 \Leftrightarrow \frac{1}{2} \int d^3x (i\hbar(\partial_t \Psi^*) \Psi - i\hbar \Psi^* (\partial_t \Psi)) = \hbar\omega. \quad (13)$$

Por otra parte, la condición de cuantización de Einstein de los modos ondulatorios y monocromáticos del campo electromagnético se basa en la relación

$$\frac{1}{8\pi} \int d^3x (\vec{E}^* \cdot \vec{E} + \vec{B}^* \cdot \vec{B}) = \hbar\omega \Leftrightarrow \frac{1}{8\pi c} \int d^3x (i\hbar(\partial_t E_s^*)A_s - i\hbar E_s^*(\partial_t A_s)) = \hbar\omega. \quad (14)$$

La similitud entre las Ecs.(13-14) es evidente; la robustez de estas expresiones proviene de que unidas a ellas hay relaciones de conservación para densidades de probabilidad (ondas materiales) y de energía (campo electromagnético). Adicionalmente este enfoque da pie a incorporar en la teoría *a la* Dirac la creación y aniquilación de partículas materiales y fotones con características predeterminadas por la estructura espacial dada por ψ , o espacial y de polarización dada por A_μ . Para ello las condiciones de normalización se extienden a condiciones de ortonormalización

$$\int d^3x \psi_a^* \psi_b = \delta_{ab}, \quad \frac{1}{8\pi} \int d^3x (\vec{E}_a^* \cdot \vec{E}_b + \vec{B}_a^* \cdot \vec{B}_b) = \hbar\omega \delta_{ab}. \quad (15)$$

Nótese que, aprovechando la expresión 11, podríamos considerar para el campo E. M. un producto escalar equivalente

$$\frac{1}{4\pi} \int d^3x E_s^* \overleftrightarrow{\partial}_{ct} A_s = \hbar\omega \delta_{ab}, \quad \overleftrightarrow{\partial}_{ct} = \frac{1}{2} [\overleftarrow{\partial}_{ct} - \overrightarrow{\partial}_{ct}] \quad (16)$$

Para ilustrar la relevancia interpretativa de las Ecs. (11-12) consideremos al momento angular del campo electromagnético en una región \mathcal{V} alrededor de un punto \vec{r}_0 ,

$$\vec{J} = \frac{1}{4\pi} \int_{\mathcal{V}} d^3x (\vec{r} - \vec{r}_0) \times [\vec{E}(\vec{r}, t) \times \vec{B}(\vec{r}, t)] \quad (17)$$

¿Puede descomponerse este momento como constituido por un momento angular orbital $\vec{L}(\vec{r}_0)$ y otro de espín \vec{S} ? Esta pregunta la plantearon los fundadores de la física cuántica (incluida la óptica cuántica), Darwin[2], Humblet[3] y de Broglie[4] concluyendo que una descomposición adecuada es[5]

$$\vec{J} = \frac{1}{4\pi} \int_{\mathcal{V}} d^3x E_s [(\vec{r} - \vec{r}_0) \times \vec{\nabla}] A_s + \frac{1}{4\pi} \int_{\mathcal{V}} d^3x \vec{E} \times \vec{A} - \frac{1}{4\pi} \int_{\mathcal{S}} (d\vec{s} \cdot \vec{E}) [(\vec{r} - \vec{r}_0) \times \vec{A}] \quad (18)$$

$$= \vec{L}(\vec{r}_0) + \vec{S} - \frac{1}{4\pi} \int_{\mathcal{S}} (d\vec{s} \cdot \vec{E}) [(\vec{r} - \vec{r}_0) \times \vec{A}] \quad (19)$$

siempre y cuando A_μ satisfaga la norma de Lorentz. Si consideramos ondas electromagnéticas descritas en términos de la parte real de un campo eléctrico complejo \vec{E} , y recordamos que en la norma de Coulomb $\vec{E} = -\partial_{ct} \vec{A}$, al escribir

$$\vec{L}(\vec{r}_0) = \frac{1}{4\pi} \int_{\mathcal{V}} d^3x E_s^* [(\vec{r} - \vec{r}_0) \times \vec{\nabla}] A_s \quad (20)$$

se observa que esta propiedad depende directamente de la fase del campo. Simultáneamente \vec{S} será dependiente de la dirección local del campo eléctrico, es decir de la polarización.

En términos del teorema de Noether, la expresión de \vec{J} deriva de efectuar una rotación espacial sobre un espacio isotropo. Al ser el campo electromagnético un campo vectorial descrito por el potencial electromagnético A_μ , una rotación infinitesimal por un ángulo $\delta\varphi$ se producen combinaciones lineales de las componentes de \vec{A} . Es decir, para cada componente A_r

$$A'_r = A_r + \delta\varphi \sum_{s=1,2,3} \widehat{M}_{rsij} \phi_s, \quad (21)$$

$$= A_r + \delta\varphi \left[\widehat{L}_{ij} \phi_r + \sum_{s=1,2,3} \widehat{S}_{rsij} \phi_s \right], \quad (22)$$

$$\widehat{L}_{ij} = -i(x_i \partial_j - x_j \partial_i), \quad (23)$$

$$\widehat{S}_{rsij} = -i[\delta_{ri} \delta_{sj} - \delta_{rj} \delta_{si}], \quad (24)$$

Los índices i, j se determinan utilizando el tensor de Levi-Civita ε_{tij} pra rotaciones alrededor del eje- t . Aquí \widehat{L}_{ij} es el operador de momento angular orbital y \widehat{S}_{rsij} es el operador tensorial para espín 1.

Esto puede entenderse mejor considerando un ejemplo sencillo: una onda monocromática con simetría a circular cilíndrica. En este caso hay una dirección espacial destacada que corresponde a la orientación del eje del cilindro que identificamos como el eje- x_3 . Direcciones de polarización naturales son las circulares cilíndricas $\hat{e}_1 \pm i\hat{e}_2$. Los modos naturales del campo se les conoce como haces Bessel. Ellos cuentan con una proyección de momento angular total bien definida con valores $m \pm 1$. El término m proviene de factores de fase $e^{im\varphi}$ donde φ es el ángulo azimutal alrededor del eje- x_3 . El término ± 1 esta relacionado con la presencia de alguno de los vectores $\hat{e}_1 \pm i\hat{e}_2$ en la expresión del campo eléctrico del modo en cuestión [6]. Hay que destacar que dada una función cualquiera de la coordenada radial ρ que localmente este multiplicada por un factor $e^{im\varphi}$, esta función no está bien definida para $\rho = 0$, *i. e.*, hay una singularidad de fase en ese punto a la que en el contexto en que estamos trabajando se le conoce como vórtice óptico. Su estudio es competencia de la llamada Óptica Singular.

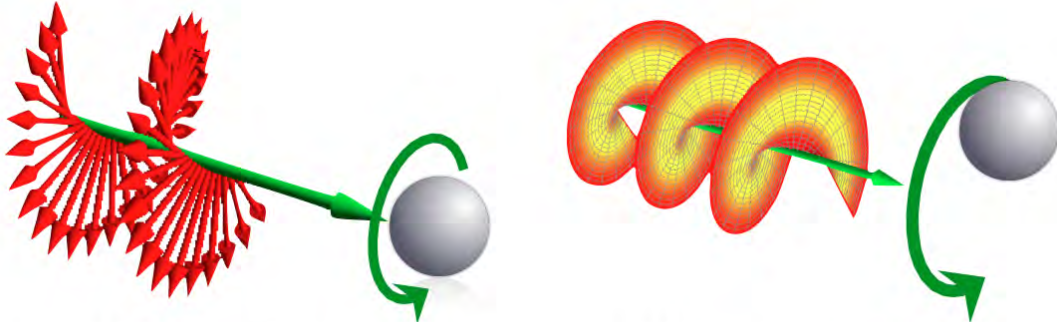


Figure 1: El momento angular de espín se asocia con la variación ante propagación de la dirección del campo eléctrico en una onda electromagnética mientras que el momento angular orbital es debido a variaciones en la fase. Ha sido comprobado experimentalmente que estos momentos pueden inducir sobre partículas microscópicas rotaciones sobre su propio eje y rotaciones alrededor de un eje externo respectivamente.

Esto podría parecer un estudio de interés meramente académico. No es así debido a los avances técnicos y tecnológicos recientes que permiten generar haces de luz con estructura de fase y polarización bajo diseño con relativa facilidad. El ejemplo paradigmático corresponde a la generación de luz con momento angular orbital en el régimen paraxial. La aproximación paraxial se utiliza la descripción de sistemas ópticos cuando la luz puede describirse como conformada por rayos que forman ángulos pequeños con respecto a un eje óptico. En esta aproximación y en la región donde la intensidad del modo es apreciable, el campo posee una polarización global; la amplitud y fase del campo quedan aproximadamente determinadas por una función escalar. La distribución de intensidad y fase de la luz proveniente de un láser suele cumplir esta aproximación. En 1992, Allen *et al* [7] encontraron que la luz láser con una distribución de amplitud Laguerre-Gaussiana (descrita por un polinomio de Laguerre superpuesto a una modulación gaussiana) tiene un momento angular orbital bien definido, pues su dependencia en el ángulo azimutal es de la forma $e^{im\varphi}$. Utilizaron un sistema óptico sencillo para su generación. Un punto muy importante es que propusieron un experimento para medir el par mecánico inducido por la transferencia del momento angular orbital asociado con dicha transformación. Esto es análogo a un experimento ejecutado por Beth en 1936 [8] en que se mostraron los efectos mecánicos del momento angular de espín de la luz al inducir rotaciones de un cristal birrefringente.

Un punto a destacar es que a nivel atómico se ha comprobado la conservación de energía, momento lineal y momento angular en las interacciones de las componentes del átomo con el campo E.M. Así la conservación de energía lleva a que el color de la luz emitida o absorbida por un átomo esta determinada por la diferencia energética de los estados internos del átomo involucrado, $\hbar\omega \approx E_n - E_m$; los efectos de rebote que se utilizan para enfriar los átomos en la técnica Doppler[9, 10] son debidos a la conservación de momento lineal en las transiciones, y las reglas de selección dipolares eléctricas se deben a la conservación de momento angular de polarización que modifica el momento angular de los estados internos atómicos.

3 Modos vectoriales dentro de un espejo paraboloidal ideal y sus simetrías.

El espejo paraboloidal es un dispositivo óptico cuya sección transversal corresponde a una paraboloides de revolución. Rayos paralelos entrantes se reflejan todos en el mismo punto focal, lo cual no ocurre para un espejo con forma esférica, que presenta las llamadas aberraciones esféricas. La óptica de rayos sugiere correctamente que los espejos parabólicos son adecuados para enfocar con precisión los haces colimados, tanto dentro como fuera de la aproximación paraxial. Sin embargo, esa descripción no proporciona información de las propiedades de fase y de polarización de cada uno de los modos que constituyen el conjunto completo de modos del campo electromagnético dentro del espejo. En los próximos párrafos mostraremos como las simetrías geométricas de estos espejos pueden utilizarse para identificar a estos modos vectoriales, establecer en forma compacta la ecuación que discretiza los valores de las variables que caracterizan a cada modo, y relacionar estas variables con propiedades mecánicas del campo electromagnético que describen [11].

La ecuación de onda monocromática en coordenadas paraboloidales

$$\nabla^2 \Psi(\zeta, \eta, \varphi) = \frac{\omega^2}{c^2} \Psi(\zeta, \eta, \varphi) \quad (25)$$

$$= \frac{4}{\zeta + \eta} \left[\frac{\partial}{\partial \zeta} \zeta \frac{\partial}{\partial \zeta} + \frac{\partial}{\partial \eta} \eta \frac{\partial}{\partial \eta} \right] \Psi + \frac{1}{\zeta \eta} \frac{\partial^2 \Psi}{\partial \varphi^2} \quad (26)$$

es separable. En la región interna de un paraboloides dado, la expresión para cualquier onda escalar se puede escribir como combinación lineal de los modos elementales

$$\Psi(\zeta, \eta, \varphi) = \sum_{\kappa, m} c_{\kappa, m} e^{im\varphi} V_{\kappa, |m|}(\omega\zeta/c) V_{-\kappa, |m|}(\omega\eta/c), \quad (27)$$

$$V_{\kappa, m}(z) = z^{|m|/2} e^{-iz/2} M\left(\frac{|m|+1}{2} - i\kappa, |m|+1; iz\right). \quad (28)$$

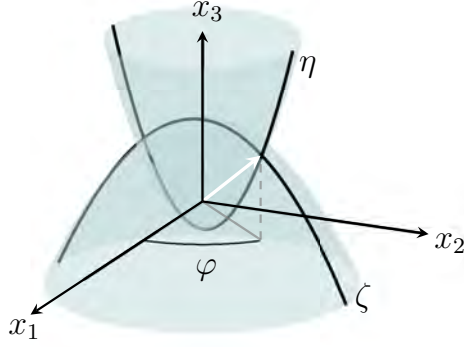


Figure 2: Coordenadas paraboloidales $\{\zeta, \eta, \varphi\}$. Las superficies con ζ (η) constante corresponden a paraboloides de revolución alrededor del eje x_3 , con foco de las parábolas localizado en el origen, y que abren hacia abajo (arriba); la tercera coordenada φ corresponde al ángulo azimutal.

Las funciones de Whittaker $M_{i\kappa, \mu}$ son útiles en diversas áreas de la física destacando el caso de la física atómica y nuclear. La transformada de Fourier de $\Psi(\zeta, \eta, \varphi)$ involucra funciones simples

$$e^{im\varphi} V_{\kappa, |m|}(\omega\zeta/c) V_{-\kappa, |m|}(\omega\eta/c) \equiv a_{\kappa, m} \int_{\mathbb{S}_2^{(\mathbf{k})}} e^{i(\omega/c)(\hat{\mathbf{k}} \cdot \mathbf{r})} f_{\kappa, m}(\theta_k, \varphi_k) d\Omega_k \quad (29)$$

$$f_{\kappa, m}(\theta_k, \varphi_k) = \frac{[\tan(\theta_k/2)]^{-i2\kappa} e^{im\varphi_k}}{\sin \theta_k} \frac{1}{2\pi}, \quad (30)$$

$$a_{\kappa, m} = (-i)^{|m|} \frac{|\Gamma(|m| + 1)|^2}{\Gamma(\frac{|m|+1}{2} + i\kappa) \Gamma(\frac{|m|+1}{2} - i\kappa)}, \quad (31)$$

Las simetrías detrás de la separabilidad de la ecuación de onda en coordenadas paraboloidales [12] son inducidas por el generador de rotaciones a lo largo del eje x_3 , que sabemos que para un campo escalar esta dado por

$$\hat{L}_3 = (\hat{\mathbf{r}} \times \hat{\mathbf{P}})_3, \quad \hat{\mathbf{r}} = \vec{r}, \quad \hat{\mathbf{P}} = -i\vec{\nabla} \quad (32)$$

y el operador parabólico

$$\frac{1}{2} [\{\hat{L}_1, \hat{P}_2\} - \{\hat{L}_2, \hat{P}_1\}], \quad (33)$$

que resulta de la resta del producto simetrizado del generador de rotaciones a lo largo del eje x_1 y traslaciones a lo largo del eje x_2 , y el producto simetrizado del generador de rotaciones a lo largo del eje x_2 y las traslaciones a lo largo del eje x_1 . El último operador es la tercer componente del operador obtenido del producto de los operadores de momento angular y lineal, $(\hat{\mathbf{L}} \times \hat{\mathbf{P}} - \hat{\mathbf{P}} \times \hat{\mathbf{L}})/2$. Es la parte cinética del bien estudiado vector de Runge-Lenz, que a su vez es el generador de una peculiar simetría asociada a una partícula cargada en presencia del potencial de Coulomb; su conexión con una descripción parabólica de tal sistema se hace explícita en Ref. [13].

Los generadores de simetría paraboloidal actuando sobre cada componente s de un campo vectorial son

$$\hat{J}_3 \phi_s = -i(x_1 \partial_2 - x_2 \partial_1) \phi_s + \sum_{j=1,2,3} \hat{S}_{sj12} \phi_s, \quad (34)$$

$$\hat{\mathfrak{A}}_3 \phi_s = \frac{1}{2} [\{\hat{M}_{sj23}, \hat{P}_2\} - \{\hat{M}_{sj31}, \hat{P}_1\}] \phi_s. \quad (35)$$

Las expresiones para la densidad de energía, Ec. 11, momento lineal, Ec. 12, y momento angular, Ec. 18, sugieren que la descripción natural de los campos con una simetría dada se obtiene si estos son eigenfunciones de los generadores de esas simetrías. En la Ref. [11] se muestra como garantizar que los campos ondulatorios \vec{E} y \vec{B} satisfagan

$$\hat{J}_3 \vec{E}_{m_j, \kappa} = m_j \vec{E}_{m_j, \kappa}, \quad \hat{J}_3 \vec{B}_{m_j, \kappa} = m_j \vec{B}_{m_j, \kappa} \quad (36)$$

$$\hat{\mathfrak{A}}_3 \vec{E}_{m_j, \kappa} = 2\kappa \vec{E}_{m_j, \kappa}, \quad \hat{\mathfrak{A}}_3 \vec{B}_{m_j, \kappa} = 2\kappa \vec{B}_{m_j, \kappa}. \quad (37)$$

en el caso paraboloidal. En ese artículo se trabaja en la norma de Coulomb y el campo \vec{A} se escribe ya sea como el rotacional de un potencial $\vec{\pi}$ (los llamados modos eléctricos) o como el rotacional de un rotacional de un potencial π (modos magnéticos). Estas dos opciones corresponden a las dos opciones independientes en polarización para la luz, solo que ahora hay una dependencia local en

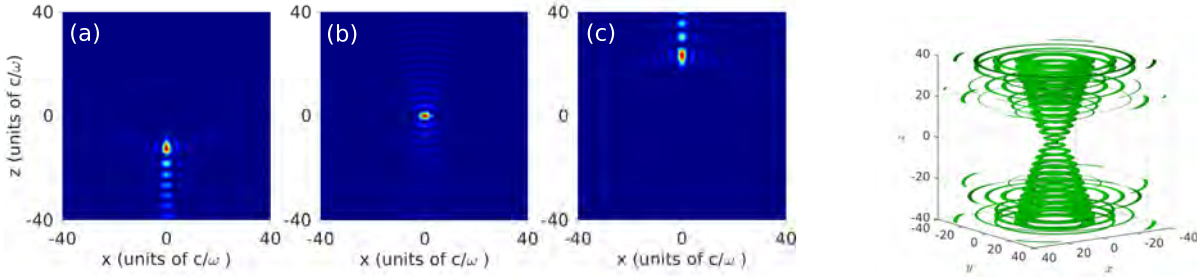


Figure 3: Ilustración de la intensidad del campo eléctrico para diversos modos paraboloidales. En los casos (a-c) se trata de modos con $\kappa > 0$, $\kappa = 0$ y $\kappa < 0$ y ejemplifican como κ no solo fija el valor del momento paraboloidal sino además determina el plano con x_3 constante en el que la intensidad de \vec{E} es máxima. La última figura ilustra las curvas de isointensidad del modo eléctrico con $\kappa = 0$. Este modo es el que está altamente enfocado al origen y que aún así presenta isosuperficies con intensidades cercanas al cincuenta por ciento como las aquí ilustradas.

la orientación del campo eléctrico correspondiente. Los resultados establecen relaciones que deben cumplir los coeficientes $c_{\kappa; m}$ de la Eq. (27) que especifican cada una de las componentes de $\vec{\pi}$ que adicionalmente son solución de la ecuación de onda. Podríamos decir entonces que estos modos tienen bien definidas tanto la componente del momento angular como la la componente del vector de Runge-Lenz total—que hace las veces de momento paraboloidal— ambas a lo largo del eje- x_3 .

Ahora bien, sobre la superficie de un espejo paraboloidal ideal definida por la condición $\zeta = \zeta_0$, se debe cumplir que

$$\vec{B} \cdot \hat{e}_\zeta \Big|_{\zeta=\zeta_0} = 0, \quad \vec{E} \cdot \hat{e}_\eta \Big|_{\zeta=\zeta_0} = 0, \quad \vec{E} \cdot \hat{e}_\varphi \Big|_{\zeta=\zeta_0} = 0. \quad (38)$$

Se puede mostrar que los valores de ese momento paraboloidal quedan restringidos con la ecuación

$$\mathcal{W}_{\kappa, m_j}^2(\omega\zeta_0/c) = -\frac{d_+}{d_-}, \quad d_\pm = \frac{1}{2} \pm i \frac{\kappa}{m_j}, \quad \mathcal{W}_{\kappa, m_j} = \frac{V_{\kappa+i/2, m_j}(\omega\zeta_0/c)}{V_{\kappa-i/2, m_j}(\omega\zeta_0/c)}. \quad (39)$$

Dado que tenemos expresiones explícitas para $(\vec{E}; \vec{B})$ de los modos parabólicos E.M. podemos estudiar en detalle su configuración. Una primera observación es que *todos* los modos están localizados espacialmente. Adicionalmente su intensidad máxima está ubicada en la cercanía de un plano perpendicular al eje Z y que depende del valor de κ . Conforme el valor de m_j aumenta estas regiones de localización forman anillos de radio cada vez mayor. La polarización y fase de los campos eléctricos y magnéticos presenta una topología rica y no trivial. Se encuentran regiones donde el campo eléctrico y magnético tienen muy diferente magnitud, singularidades de fase, vórtices vectoriales y fuertes gradientes de intensidad. Estas propiedades deben estudiarse en profundidad en términos de su efecto en la interacción con los sistemas atómicos [14].

4 Comentarios finales.

En este escrito se habló con relativo detalle de las propiedades mecánicas asociadas a sistemas con simetría Cartesiana, circular cilíndrica y paraboloidal. Otros ejemplos presentados durante la charla correspondieron a haces con simetrías elíptico cilíndrica y parabólico cilíndrica; para ellos además de poder desarrollarse el formalismo análogo al aquí descrito se han realizado experimentos con átomos ultrafríos y con gases atómicos térmicos para comprobar: (i) la transferencia de momento angular elíptico de la luz en procesos de óptica no lineal [15] y (ii) la transferencia de momento paraboloidal a la materia [16]. Invito al lector interesado a leer los artículos mencionados y en su caso, investigar a las variables dinámicas de la luz y sus efectos sobre la materia.

References

- [1] C. Itzykson and J. B. Zuber, *Quantum Field Theory*, (MacGraw-Hill NY, 1980) Chap. I, p. 8.
- [2] C. G. Darwin, *Proc. R. Soc. London, Ser. A* **136**, 36 (1932).
- [3] J. Humblet, *Physica (Amsterdam)* **10**, 585 (1943).
- [4] L. de Broglie, *Mécanique Ondulatoire du Photon et Théorie Quantique des Champs* (Gauthier-Villars, Paris, 1949), Chap. VI, p. 65.
- [5] L. Mandel and E. Wolf, *Optical Coherence in Quantum Optics*, Cambridge University Press 1995, p.p. 442.

- [6] R. Jáuregui and S. Hacyan, *Phys. Rev. A* **71**, 033411 (2005)
- [7] L. Allen, M. W. Beijersbergen, R. J. C. Spreeuw, and J. P. Woerdman, *Phys. Rev. A* **45**, 8185 (1992).
- [8] R. A. Beth, *Phys. Rev.* **50**, 115 (1936).
- [9] D. J. Wineland and H. Dehmelt, *Bulletin of the A. P. S.* **20**, 637 (1975).
- [10] T. W. Hänsch and A. L. Shawlow, *Optics Comm.* **13**, 6 (1975).
- [11] R. Gutiérrez-Jáuregui and R. Jáuregui, *Phys. Rev. A* **98**, 043808 (2018).
- [12] C. P. Boyer, E. G. Kalnins, and W. Miller Jr., *Nagoya Math. J.* **60**, 35-80 (1976).
- [13] M. Bander and C. Itzykson, "Group Theory and the Hydrogen Atom ", *Rev. Mod. Phys.* **38**, 330 (1966).
- [14] R. Gutiérrez-Jáuregui and R. Jáuregui, *Scientific Reports-Nature* **10**, 17383 (2020).
- [15] L. A. Mendoza-López, J. G. Acosta-Montes, I. F. Ángeles-Aguillón, D. Sierra-Costa, Y. M. Torres, R. Jáuregui, and D. Sahagún-Sánchez, *Phys. Rev. Research* **3**, 033170 (2021).
- [16] C. L. Hernández-Cedillo, S. Bernon, H. Hattermann, J. Fortágh, and R. Jáuregui, "Scattering of dilute thermal atom clouds on optical Weber beams", *Phys. Rev. A* **87**, 023404 (2013).

Cálculo de propiedades ópticas de metamateriales*

Merlyn Jaqueline Juárez-Gutiérrez y W. Luis Mochán
Instituto de Ciencias Físicas

3 de noviembre de 2021

Resumen

Presentamos una introducción a los metamateriales, algunas de sus propiedades ópticas y ejemplos de sus usos. Desarrollamos una teoría eficiente para el cálculo de su permitividad macroscópica en sistemas binarios o con más componentes, en el caso no retardado y en el caso general, y presentamos su implementación en un paquete computacional y su uso. Finalizamos discutiendo algunas aplicaciones del mismo para el diseño de dispositivos ópticos nanoestructurados optimizados y discutimos las propiedades lineales y no lineales obtenidas.

1. Introducción

Un metamaterial es un material artificial formado por dos o más materiales alternados. Los metamateriales se pueden definir como *un arreglo de elementos estructurales artificiales, diseñados para alcanzar propiedades electromagnéticas ventajosas e inusuales*[1], de acuerdo al *Virtual Institute for Artificial Electromagnetic Materials and Meta-Materials*. Dichas propiedades están determinadas por sus constituyentes básicos, a los que se denomina ocasionalmente como *meta-átomos*, los cuales son objetos hechos de materiales usuales, así como por su forma y disposición, y pueden ser muy distintas a las de los materiales que los conforman, llegando a a ser muy exóticas. Pueden ser diseñadas y entonadas escogiendo las formas, estructuras internas, tamaños, orientaciones mutuas, etc., de sus meta-átomos. Sus funciones repuesta pueden ser modificadas mediante señales externas e internas y pueden ser controladas mediante microprocesadores programables.[2]

1.1. Materiales plasmónicos

Si un metamaterial tiene componentes metálicos, estos pueden presentar resonancias asociadas a los movimientos oscilatorios colectivos de sus electrones de conducción, denominados de acuerdo a sus características como plasmones de bulto, de superficie o localizados. La frecuencia de estos movimientos en un sistema infinito se

*Curso presentado en la XVIII Escuela de Verano en Física, ICF- e IF-UNAM, Cuernavaca, Mor. y Cd de México, junio 21-julio 2, 2021

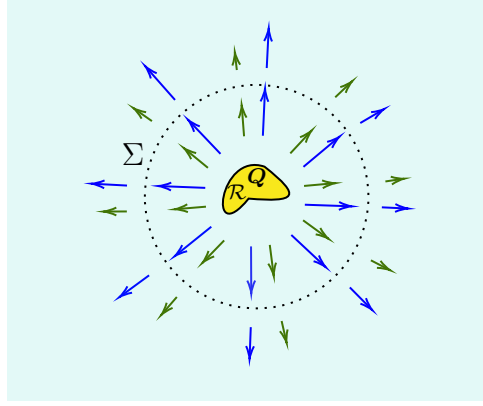


Figura 1: Acumulación de carga Q en cierta región \mathcal{R} dentro de un conductor homogéneo, rodeada por una superficie Σ sobre la cual se aplica la ley de Gauss al campo producido por Q , campo que produce una densidad de corriente que fluye a través de Σ . La corriente modifica Q conforme transcurre el tiempo t produciendo oscilaciones de plasma.

denomina como frecuencia de plasma ω_p . Para estimarla, considere el modelo más simple de un conductor, un gas de electrones que en equilibrio tienen una densidad de número n_0 , y que son libres de moverse en un entorno positivamente cargado, de manera que el sistema en equilibrio sea neutro. Si debido a alguna compresión o rarefacción del gas de electrones se produjera una acumulación de carga Q localizada en alguna región \mathcal{R} , ésta produciría un campo eléctrico $\mathbf{E}(\mathbf{r})$, como ilustra la figura 1. De acuerdo a la segunda ley de Newton, los electrones adquirirían una aceleración $d^2\mathbf{r}/dt^2 = -e\mathbf{E}(\mathbf{r}, t)/m$, donde m y $-e$ son la masa y la carga eléctrica electrónica. La velocidad adquirida por los electrones resultaría en una corriente eléctrica $\mathbf{j}(\mathbf{r}, t) = -n_0e\mathbf{v} = -n_0e(d\mathbf{r}/dt)$ que obedece la ecuación de movimiento por $\partial\mathbf{j}(\mathbf{r}, t)/\partial t = -\frac{n_0e^2}{m}\mathbf{E}(\mathbf{r}, t)$. Integrando la ecuación diferencial de la corriente sobre una superficie Σ que rodee completamente la carga y usando la ecuación de continuidad y la Ley de Gauss para el campo eléctrico obtenemos una ecuación diferencial para la carga Q encerrada en Σ ,

$$\frac{d^2Q}{dt^2} = -\frac{4\pi n_0 e^2}{m}Q, \quad (1)$$

la cual es una ecuación diferencial idéntica a la de un oscilador armónico simple como el que se ilustra en la fig. 2: Sustituyendo la Ley de Hooke $F = -ky$ en la segunda ley de Newton $ma = F$ para un oscilador con constante k y masa m , obtenemos

$$\frac{d^2y}{dt^2} = -\omega^2 y, \quad (2)$$

donde $\omega = \sqrt{k/m}$, y cuya solución, $y(t) = y_0 \cos(\omega t + \phi_0)$ es un movimiento periódico con una frecuencia ω que depende de k y m . Comparando las ecs. y (1) y (2) notamos que puede existir carga en el seno de nuestro metal modelo, pero ésta oscilaría con la *frecuencia de plasma* ω_p dada por

$$\omega_p^2 = \frac{4\pi n_0 e^2}{m}. \quad (3)$$

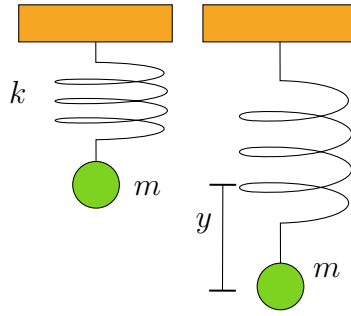


Figura 2: Oscilador armónico de masa m y constante k en su posición de equilibrio y desplazado una distancia y .

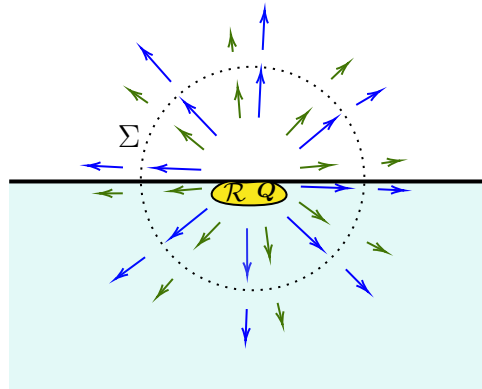


Figura 3: Región \mathcal{R} en la superficie de un conductor semi-infinito en la que hay una carga Q , la cual produce un campo eléctrico $\mathbf{E}(\mathbf{r}, t)$ (líneas azules) y éste a su vez produce una densidad de corriente (líneas verdes) que atraviesan aquella parte de la superficie lejana Σ que se halla dentro del conductor.

La repulsión mutua entre electrones los impulsa lejos de regiones en que haya una densidad electrónica excedente, por arriba de su valor nominal. El movimiento consecuente prosigue aun después de que el sistema se neutraliza debido a la inercia electrónica, que los hace proseguir su camino hasta que en la región original disminuye tanto la densidad de electrones que aparezca una carga neta positiva que frena a los electrones en fuga y los hace regresar, hasta que su repulsión mutua los frena, habiendo regresado a la configuración inicial. Este proceso se repite periódicamente y su frecuencia ω_p está relacionada con la repulsión coulombiana, proporcional a e^2 , la densidad de de número electrónica n_0 y la inercia electrónica caracterizada por m .

En lugar de un medio infinito, consideremos ahora un medio semiinfinito separado del vacío por una superficie plana. Un análisis análogo nos permite obtener la frecuencia del *plasmón de superficie*, considerando ahora un exceso de carga Q en una región R en la interfaz, como ilustra la fig. 3, el cual produce un campo eléctrico $\mathbf{E}(\mathbf{r}, t)$ que induce corrientes en el conductor.

Usando la ecuación dinámica de la densidad de corriente podemos escribir la ecuación dinámica de la carga como hicimos en el caso del plasmón de bulto, con la diferencia que la densidad de corriente en este caso sólo fluye a través de la mitad

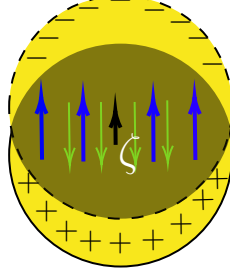


Figura 4: Esfera metálica cuyos electrones han sido desplazados en la dirección de ζ (flecha negra), dejando una ausencia de carga en la dirección opuesta induciendo polarización en el sistema \mathbf{P} (flechas verdes) y una carga superficial la cual produce un campo eléctrico $\mathbf{E}(\mathbf{r}, t)$ (flechas azules).

$\Sigma/2$ de la superficie Σ que se halla en el interior del metal,

$$\begin{aligned}
 \frac{d^2}{dt^2}Q &= - \int_{\Sigma} d\mathbf{a} \cdot \frac{\partial}{\partial t} \mathbf{j}, \\
 &= ne \int_{\Sigma/2} d\mathbf{a} \cdot \frac{\partial}{\partial t} \mathbf{v}, \\
 &= \frac{1}{2} \frac{ne^2}{m} \int_{\Sigma} d\mathbf{a} \cdot \mathbf{E}, \\
 &= - \frac{2\pi ne^2}{m} Q.
 \end{aligned} \tag{4}$$

Comparando esta ecuación con la ec. (2) identificamos la frecuencia del *plasmón de superficie* ω_{sp} , dada por

$$\omega_{sp}^2 = \frac{2\pi ne^2}{m} = \frac{\omega_p^2}{2}. \tag{5}$$

En lugar de un sistema semiinfinito, consideremos ahora un sistema finito consistente en una partícula metálica separada del vacío por una superficie esférica. Supongamos que perturbamos esta esfera moviendo todos sus electrones una separación ζ respecto a su posición de equilibrio, lo cual induce una polarización $\mathbf{P} = -n_0 e \zeta$, como ilustra la fig. 4, El desplazamiento de los electrones hacia uno de los hemisferios de la esfera genera un exceso de carga negativa en su superficie y un exceso de carga positiva en el otro, descrita por la densidad de carga superficial $\sigma = \mathbf{P} \cdot \hat{n}$, donde \hat{n} es un vector unitario radial. Estas cargas producen un campo eléctrico $\mathbf{E}(\mathbf{r}, t) = -(4\pi/3)\mathbf{P}$ donde empleamos el *factor de despolarización* $-4\pi/3$ de una esfera. Este campo acelera las cargas de acuerdo a $m d^2 \zeta / dt^2 = -e \mathbf{E}(\mathbf{r}, t)$. Escribiendo al campo en términos de ζ , $\mathbf{E} = -(4\pi n_0 e^2 / 3m) \zeta$, la ecuación de movimiento para ζ se convierte en,

$$\frac{d^2 \zeta}{dt^2} = - \frac{4\pi n_0 e^2}{3m} \zeta \tag{6}$$

de la cual obtenemos la frecuencia ω_d de las oscilaciones de carga en una esfera, a las que se les denominan como *plasmón dipolar*,

$$\omega_d^2 = \frac{4\pi ne^2}{3m} = \frac{\omega_p^2}{3}. \tag{7}$$



Figura 5: Copa de Licurgo, fabricada en la Roma tardía del siglo IV D.C., cuya fabricación resultante muestra propiedades ópticas excepcionales. Es una taza para beber de vidrio que se ve verde o roja, dependiendo de cómo es iluminada. ©The Trustees of the British Museum.

Estos ejemplos muestran que en un metal los electrones pueden animarse de movimientos colectivos asociados a ciertas frecuencias de resonancia, las cuales a su vez dependen de la geometría, como ilustramos estudiando el caso de un sistema infinito, uno semiinfinito y una esfera. Más aún, si colocamos las partículas metálicas en el seno de una matriz dieléctrica habría un corrimiento adicional en su frecuencia de resonancia debido a las cargas inducidas en la superficie del dieléctrico. Si además hubiese un gran número de esferas, sus interacciones mutuas a través de los campos electromagnéticos inducidos podrían generar corrimientos adicionales de las resonancias.

Un ejemplo excepcional de las propiedades que emergen al generar *metamateriales* es la copa de Licurgo, una copa de cristal tallada en la época romana tardía, decorada con un friso que muestra escenas del mito del Rey Licurgo. La copa que data del siglo IV, D.C., se produjo a partir de una pieza en bruto de vidrio soplado de unos 15mm de espesor. Las figuras se cortaron y rectificaron y se unieron a la pared del recipiente mediante pequeños puentes de vidrio. Aparte del trabajo artístico realizado en la decoración, la copa es de gran interés por las propiedades ópticas que muestra. El vidrio se ve de un color rojo-vino profundo cuando la luz lo atraviesa y de un color verde opaco cuando la luz que llega a nuestros ojos es reflejada por su superficie, como muestra la fig. 5. A este fenómeno se le denomina *dicroísmo*, y de los artefactos de vidrio romano encontrados, la copa es la que muestra dicho efecto más intensamente. [5] Estudios de la composición del vidrio muestran que tales propiedades son causadas por la presencia de finas partículas de oro, probablemente en una aleación con plata, dispersadas. Con estudios de microscopía de transmisión de electrones, TEM, por sus siglas en inglés, se pudieron determinar tamaños de las partículas de $\approx 10nm$. Se ha encontrado que contiene además partículas de diferentes metales y de materiales no metálicos. El color se debe al espectro de reflexión y de

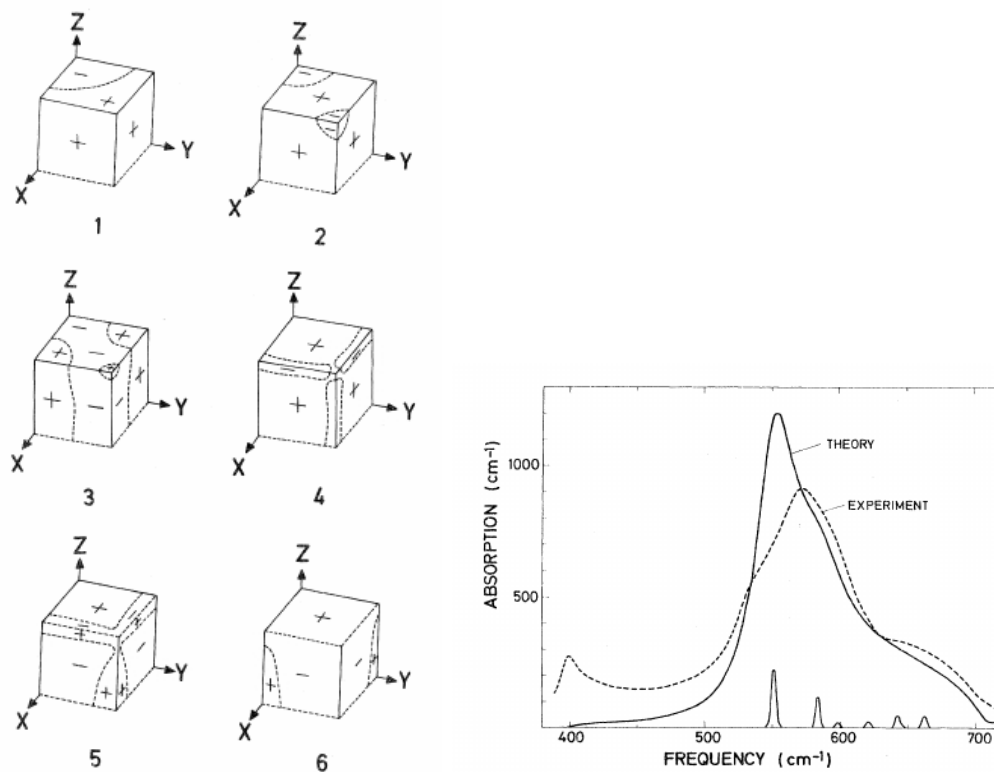


Figura 6: Modos resonantes de un cristal de sal y espectro de absorción teórico y experimental. Los signos del lado izquierdo indican la polarización en una de las ocho esquinas del cubo de sal correspondiente para cada uno de los seis modos principales. La posición y peso de cada resonancia se ilustran con las pequeñas gaussianas en la base del panel derecho, las cuales generan el espectro teórico al escalarse, ensancharse y sumarse.

transmisión del medio compuesto formado por vidrio y por las partículas metálicas. Aunque el oro es amarillo, las partículas de oro embebidas en una matriz de vidrio e interaccionando entre sí producen un color rojo.

1.2. Otras geometrías

Si consideráramos partículas con otras geometrías habría otras resonancias asociadas a la excitación de modos con patrones varios de distribución de carga. Por ejemplo, en la figura 6 mostramos resultados experimentales y los primeros resultados teóricos para los modos electromagnéticos esperados en pequeños cubos de sal [6], sus frecuencias de resonancia y su distribución asociada de carga superficial. En este caso se encontraron en lugar de un modo dipolar, como vimos para el caso de la esfera, seis modos principales y unos modos adicionales con poca fuerza de oscilador, con una polarización cuya distribución espacial muestra bastante riqueza.

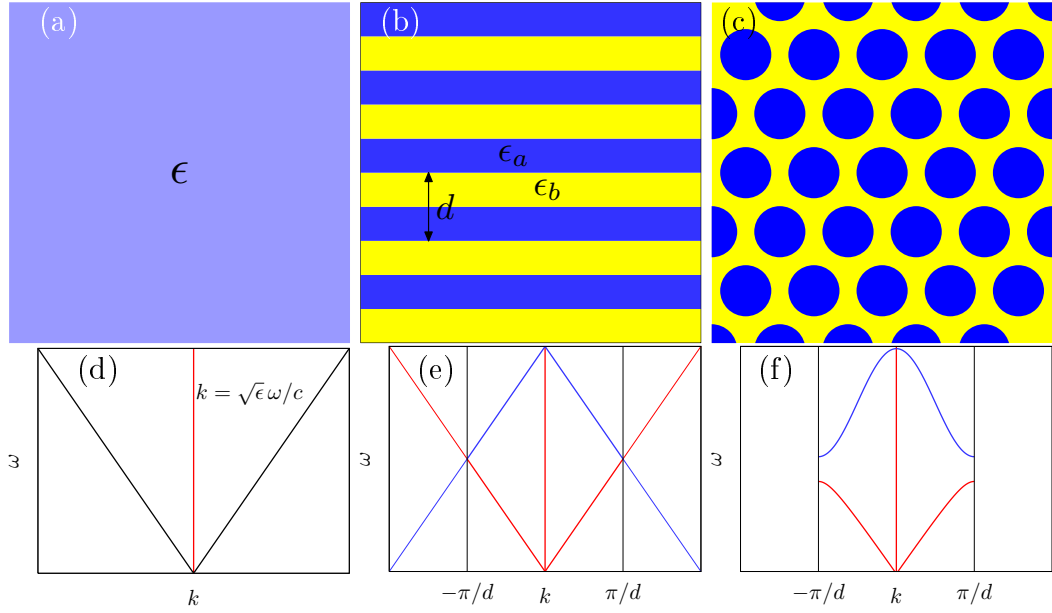


Figura 7: (a) Medio dieléctrico homogéneo con respuesta ϵ , (b) cristal fotónico unidimensional con funciones respuesta ϵ_a y ϵ_b y periodo d , y (c) cristal fotónico con periodicidad en más dimensiones. (d) Relación de dispersión ω vs k de la luz del medio homogéneo. (e) Relaciones de dispersión trasladadas por $0, \pm 2\pi/d$ en el espacio recíproco mostrando cruzamientos en el borde de la zona de Brillouin $k = \pm\pi/d$. (f) El acoplamiento abre una brecha en los puntos de degeneración evitando el cruce y formando brechas prohibidas.

1.3. Cristales fotónicos

Consideremos ahora un dieléctrico transparente no dispersivo homogéneo, como en la figura 7. La relación de dispersión de la luz en este medio está dada por $k^2 = \epsilon\omega^2/c^2$ que corresponde a las dos rectas mostradas en la (fig. 7d). Si en vez de un dieléctrico homogéneo tuviéramos un cristal artificial formado por películas de dos materiales alternados (fig. 7b), el ímpetu y el vector de onda ya no serían cantidades conservadas. Las reflexiones múltiples en las interfaces producirían ondas esparcidas en que el vector de onda cambiaría $k \rightarrow k + 2\pi n/d$ para enteros positivos y negativos n (fig. 7e), dando lugar a puntos de degeneración en que se cruzan las distintas réplicas de la relación de dispersión. El acoplamiento entre los campos esparcidos rompe la degeneración y abre *brechas fotónicas* evitando los cruces y dando origen a una relación de dispersión (fig. 7f) organizada en *bandas fotónicas* análogas a las bandas electrónicas que describen la propagación de electrones en sólidos cristalinos. Algo similar sucedería si la periodicidad fuese bidimensional o tridimensional (fig. 7c) en cuyo caso podrían producirse brechas omnidireccionales en las que la luz no se propaga en ninguna dirección. Las brechas fotónicas explican algunos fenómenos naturales, como la iridiscencia en los caparazones de diversos insectos y los colores de las alas de las mariposas, colores producidos no por pigmentos que absorben la luz, sino por pequeñas estructuras dieléctricas transparentes que forman cristales fotónicos con regiones de frecuencia en que la luz es fuertemente reflejada por corres-

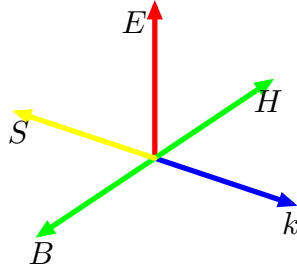


Figura 8: Relación entre las direcciones del campo eléctrico \mathbf{E} , densidad de flujo magnético \mathbf{B} , vector de onda \mathbf{k} , campo magnético \mathbf{H} y vector de Poynting \mathbf{S} en un metamaterial izquierdo isotrópico. Como $(\mathbf{E}, \mathbf{B}, \mathbf{k})$ y $(\mathbf{E}, \mathbf{H}, \mathbf{S})$ son triadas ordenadas derechas, y \mathbf{B} y \mathbf{H} son antiparalelos, entonces \mathbf{k} y \mathbf{S} son antiparalelos.

ponder a brechas en que no se puede propagar. Estos colores se llaman por su origen *colores estructurales*. Introduciendo *defectos* en cristales fotónicos pueden generarse sitios en que la luz puede ser atrapada, hecho que ha encontrado aplicaciones tales como la elaboración de fibras ópticas fotónicas.

1.4. Materiales izquierdos

Consideremos ahora un material cuya permitividad $\epsilon(\omega)$ dependa de la frecuencia y tenga un comportamiento resonante. La relación de dispersión $k^2 = \epsilon\omega^2/c^2$ implica que al pasar la resonancia, cuando ϵ adquiere valores negativos, k se vuelve imaginario y la luz no puede propagarse. Esto explica la aparición del color en los materiales comunes, en que hay frecuencias características de cada material en que absorben luz y justo arriba hay frecuencias en que no se puede propagar. Si el material tuviera además una respuesta magnética $\mu \neq 1$ la relación de dispersión cambiaría a $k^2 = \epsilon\mu\omega^2/c^2$. Si tanto ϵ como μ tuvieran resonancias cercanas, arriba de éstas podría suceder que ambas fueran negativas, $\epsilon < 0$ y $\mu < 0$. En este caso, su producto $\epsilon\mu > 0$ sería positivo y sí podría haber propagación con un vector de onda real. Sin embargo, esta propagación sería curiosa. A partir de las ecuaciones de Maxwell, por ejemplo, de las leyes de Faraday y de Gauss magnética, sabemos que para una onda plana, el campo eléctrico \mathbf{E} , la densidad de flujo magnético \mathbf{B} y el vector de onda \mathbf{k} forman una triada ordenada derecha, como ilustra la figura 8. Sin embargo, la definición del vector de Poynting y la ley de Ampère-Maxwell implican que el campo eléctrico \mathbf{E} , el campo magnético \mathbf{H} y el flujo de energía \mathbf{S} también forman una triada ordenada derecha. Sin embargo, si $\mu < 0$, entonces \mathbf{B} y \mathbf{H} apuntan en direcciones opuestas. Luego, \mathbf{S} apunta en la dirección opuesta a \mathbf{k} ! La dirección en que avanza la fase de la onda es opuesta a la dirección en que avanza la energía. Esto sólo puede ser posible si la velocidad de grupo es opuesta a la velocidad de fase. Una consecuencia curiosa de este resultado se manifiesta cuando una onda se refracta en una superficie plana. La ley de conservación del ímpetu asociada a una simetría translacional implica que las proyecciones del vector de onda \mathbf{k}_{\parallel} a lo largo de la superficie deben coincidir para las ondas reflejada, transmitida e incidente. De aquí se derivan las leyes de la reflexión y de Snell. Sin embargo, la

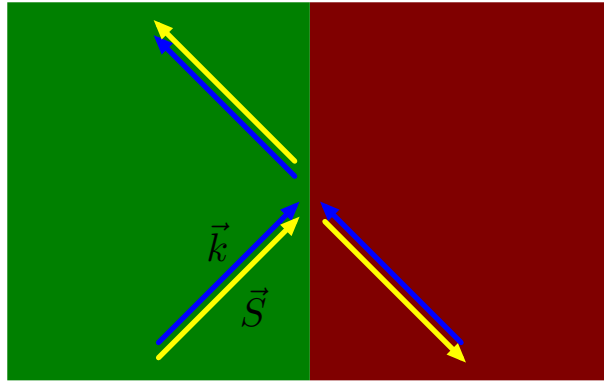


Figura 9: Una onda incide desde un medio normal (verde) sobre la interfaz que lo separa de un medio con permitividad y permeabilidad negativas (rojo). Se muestran esquemáticamente las direcciones de los vectores de onda (flechas azules) y de los vectores de Poynting (flechas amarillas) de las ondas incidente, reflejada y transmitida. La proyección sobre la interfaz de los tres vectores de onda debe coincidir. El flujo de energía de la onda transmitida debe alejarse de la interfaz. Por ello, la componente normal del vector de onda transmitido apunta hacia la interfaz.

causalidad requiere que las ondas esparcidas por la superficie, la onda incidente y la onda reflejada, deben tener un flujo de energía que se aleja de la superficie. Ello implica que cuando incide luz desde un medio ordinario hacia un medio con $\epsilon < 0$ y $\mu < 0$, la componente normal del vector de onda de la onda transmitida ¡debe apuntar hacia la superficie!, como ilustra la fig. 9. De esta figura podemos inferir que una onda que incide viajando hacia arriba se refracta hacia abajo y viceversa. Eso lleva a plantear dispositivos como el ilustrado en la fig. 10, consistente en una película plana de un metamaterial izquierdo con $\epsilon < 0$ y $\mu < 0$. La luz que emerge de una fuente puntual y viaja hacia la derecha y hacia arriba se refracta hacia abajo mientras que luz que viaja hacia abajo se refracta hacia arriba. Algo análogo sucede al emerger de la película. Es posible entonces que todos los rayos que parten de la fuente luminosa converjan en un punto, la imagen de la fuente formada por una lente

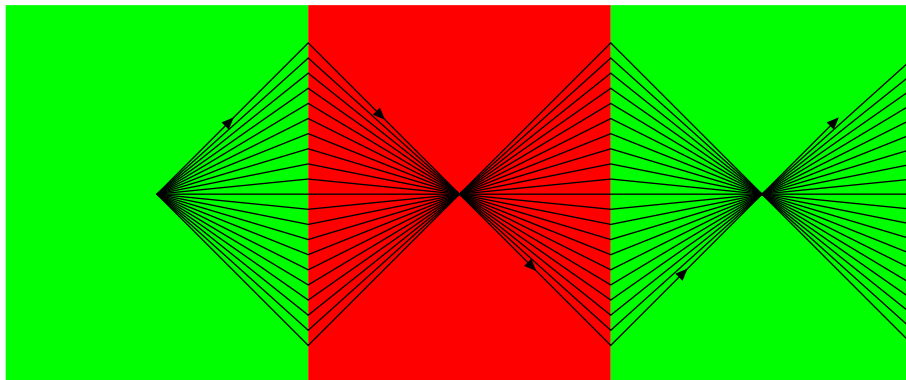


Figura 10: Lente consistente en una película plana de un metamaterial izquierdo (rojo) en el seno de un material ordinario (verde). Se ilustra la trayectoria de varios rayos que emergen de una fuente puntual de luz y convergen en un punto imagen.

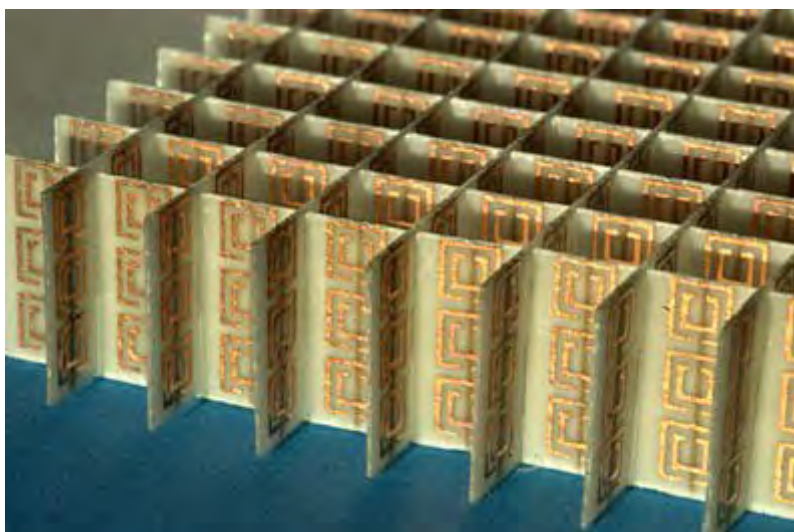


Figura 11: Metamaterial izquierdo formado por una red de parejas de anillos conductores interrumpidos (*split rings*) y pistas conductoras rectas sobre un dieléctrico.

plana.

Desafortunadamente, no existen materiales naturales en los que tanto la permitividad como la permeabilidad sean negativas a la misma frecuencia. Sin embargo, hay metamateriales artificiales que pueden describirse por una permitividad y permeabilidad efectiva que sí cumplan esta condición. La figura 11 muestra un ejemplo formado por un arreglo de parejas de anillos interrumpidos que funcionan como un circuito LC resonante. La corriente recorriendo los anillos produce un dipolo magnético, y debido a su interrupción, produce una acumulación de cargas que lo acopla con el otro anillo. Este circuito tiene una resonancia arriba de la cual la permeabilidad macroscópica es negativa. Por otro lado, una serie de pistas rectas permiten que el material se comporte en la dirección vertical como un conductor, por lo cual la permitividad es negativa abajo de la frecuencia de plasma efectiva.

1.5. Metasuperficies

Hemos visto arriba que partículas metálicas pequeñas pueden tener resonancias plasmónicas cuyas frecuencias dependen en general de su composición y de su geometría. También partículas dieléctricas pueden tener resonancias aunque estén formadas por materiales no dispersivos, siempre y cuando la longitud de onda de la luz en su interior sea conmensurable con su tamaño. Estas resonancias se deben a la interferencia constructiva entre ondas múltiplemente reflejadas por sus superficies. Por ejemplo, partículas esféricas o cilíndricas muestran *resonancias de Mie* cuando el perímetro de su sección transversal es cercano a un múltiplo de la longitud de onda. Una ventaja de estas resonancias sobre las resonancias plasmónicas para diseñar y construir dispositivos fotónicos es que las pérdidas de energía debidas a la absorción dentro del material son menores que las pérdidas óhmicas que suelen mostrar los metales. Sin embargo, estas resonancias requieren que las partículas tengan un tamaño relativamente grande conmensurable con la longitud de onda en su interior.

Sin embargo, si se emplean materiales con un índice de refracción alto, la longitud de onda dentro de estos materiales puede ser mucho menor que la correspondiente al espacio vacío, permitiendo así resonancias dieléctricas en partículas de tamaño muy pequeñas, en analogía a las resonancias plasmónicas.

Las funciones respuesta de una partícula cambian de signo conforme la frecuencia de la luz pasa de ser menor a ser mayor a su frecuencia de resonancia. Por tanto, la fase que adquiere un haz luminoso al pasar a través de una superficie cubierta por partículas depende muy sensiblemente de la cercanía de la frecuencia a la frecuencia de resonancia de las partículas, la cual a su vez, depende de la geometría. Por tanto, modulando la geometría de las partículas a lo largo de la superficie, puede modularse la fase que adquiere la luz en forma análoga a como el ancho variable de una lente o de un prisma modula la fase de los rayos de luz que los atraviesan.

La ley de Snell usual implica que a lo largo de una interfaz uniforme hay un empatamiento de fases $\phi^\alpha(\mathbf{r}_\parallel) = \mathbf{k}_\parallel^\alpha \cdot \mathbf{r}_\parallel$ entre la onda incidente, la onda reflejada y la onda transmitida, por lo cual los vectores de onda $\mathbf{k}_\parallel^\alpha$ proyectados sobre la superficie son iguales para las tres ondas $\alpha = i, r, t$. Sin embargo, si la superficie no es uniforme y a lo largo de ésta la onda transmitida y/o reflejada adquiere una fase adicional $\psi^\alpha(\mathbf{r}_\parallel)$ a la de la onda incidente, $\phi^\alpha(\mathbf{r}_\parallel) = \mathbf{k}_\parallel^i \cdot \mathbf{r}_\parallel + \psi^\alpha(\mathbf{r}_\parallel) \approx (\mathbf{k}_\parallel^i + \nabla_\parallel \psi^\alpha(0)) \cdot \mathbf{r}_\parallel$ la ley de Snell debe generalizarse,

$$\mathbf{k}_\parallel^\alpha = \mathbf{k}_\parallel^i + \nabla_\parallel \psi^\alpha, \quad (\alpha = r, t) \quad (8)$$

i.e., el ímpetu paralelo a la interfase adquiere una contribución debida a la variación de la fase adicional ψ^α . Por lo tanto, modulando la fase de una onda a lo largo de una superficie podemos manipular la dirección de la luz transmitida o reflejada. Para esto se pueden colocar partículas con un índice de refracción grande sobre una superficie ordinaria y modificar a lo largo de esta su geometría, orientación o densidad, dando lugar a una *metasuperficie*.

Como las resonancias dependen también de la polarización de la luz, este efecto puede usarse para desviar haces de luz de acuerdo a su polarización. En la fig. 12 mostramos una *metasuperficie* formada por partículas en forma de L cuya respuesta difiere cuando es iluminada con polarización circular derecha o polarización circular izquierda, y que por tanto puede separar un haz de luz no polarizada en dos haces con polarizaciones opuestas. Mediante otros arreglos se puede variar la fase a lo largo de la dirección radial, de forma de hacer converger rayos que arriben en la dirección normal sobre un punto, su foco, creando así una metalente, como la que ilustra la figura 13.

2. Teoría

Las propiedades ópticas de materiales compuestos como los presentados arriba, metamateriales, cristales fotónicos, materiales izquierdos, etc., con propiedades en ocasiones exóticas, están determinadas no sólo por su composición, sino también por su geometría. Propiedades como las relaciones de dispersión de los modos electromagnéticos que se propagan a través de un metamaterial extendido, las amplitudes de reflexión y transmisión, y relaciones de dispersión de modos electromagnéticos

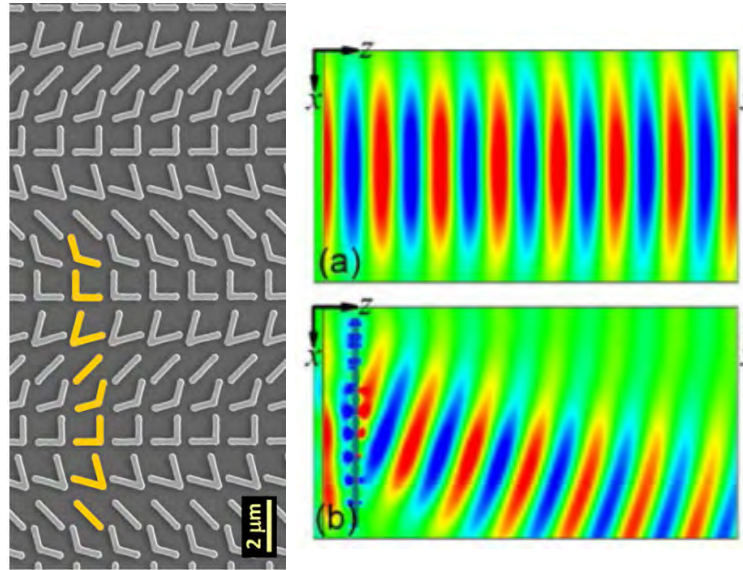


Figura 12: Metasuperficie formada por un arreglo de partículas en forma de L con ángulos variables (izquierda), capaz de separar un haz luminoso en sus componentes con polarización circular derecha e izquierda (derecha).

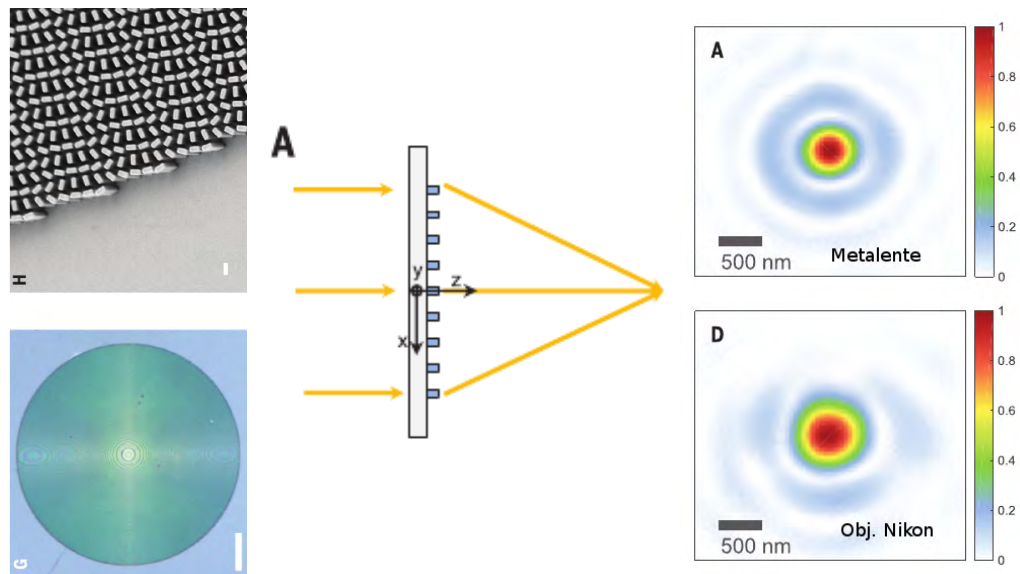


Figura 13: Metalente formada por un arreglo de prismas de alto índice de refracción con orientaciones variables (izquierda). Refracción de la luz debido a la modulación de fase al atravesar el metalente (centro). Imagen de una fuente puntual formada por la metalente y por una lente convencional (derecha).

confinados a la superficie de sistemas con fronteras, las secciones transversales de dispersión, absorción y extinción de partículas formadas por partículas, pueden ser expresadas en términos del operador dieléctrico macroscópico del compuesto a través de las soluciones de las ecuaciones de Maxwell en dicho material. En esta sección presentaremos un formalismo para obtener la respuesta macroscópica en términos de la respuesta microscópica.

2.1. Proyectores promedio y fluctuación

Primero notamos que el campo electromagnético dentro de un material inhomogéneo tiene oscilaciones relacionadas con su *textura* y la escala de variación espacial de estas oscilaciones es del orden del tamaño de las partículas que forman el compuesto así como de las distancias entre partículas vecinas. Llamamos campo macroscópico a aquel del cual hemos eliminado dichas fluctuaciones. El campo macroscópico puede tener oscilaciones espaciales, pero éstas están asociadas a las oscilaciones temporales del campo y a la longitud de onda finita de los campos que se propagan. En todo caso, es conveniente introducir dos operadores, el *promedio* $\hat{\mathcal{P}}_p$ y la *fluctuación* $\hat{\mathcal{P}}_f$, tales que al actuar sobre un campo arbitrario \mathbf{F} producen el campo promedio $\mathbf{F}_p = \hat{\mathcal{P}}_p \mathbf{F}$ y el campo fluctuante $\mathbf{F}_f = \hat{\mathcal{P}}_f \mathbf{F}$. Existen muchas formas de definir promedio. En sistemas desordenados podríamos usar el promedio de ensamble, es decir, sumar sobre N realizaciones del sistema y dividir entre N , tomando el límite $N \rightarrow \infty$. En sistemas dinámicos como un fluido, y para campos que oscilen lentamente con respecto a los tiempos característicos en que cambia el sistema, podríamos emplear un promedio temporal. Para otro tipo de sistemas podríamos tomar un promedio espacial o aplicar un filtro pasabajo en el espacio recíproco. Lo que debe ser claro es que un campo se promedia cuando se le remueven las fluctuaciones, i.e., $\hat{\mathcal{P}}_p = \hat{\mathbf{1}} - \hat{\mathcal{P}}_f$, con $\hat{\mathbf{1}}$ el operador identidad. Si pretendiéramos remover las fluctuaciones de un campo que ya hemos promediado, encontraríamos que no queda nada por remover. Esto implica que el operador promedio es idempotente $\hat{\mathcal{P}}_p^2 = \hat{\mathcal{P}}_p$, i.e., el promedio del promedio es el promedio. Análogamente, las fluctuaciones son lo que queda al remover el promedio. Por lo tanto, el operador fluctuación también es idempotente, $\hat{\mathcal{P}}_f^2 = \hat{\mathcal{P}}_f$. Finalmente, si eliminamos las fluctuaciones y el promedio, no nos queda nada, $\hat{\mathcal{P}}_p \hat{\mathcal{P}}_f = 0$, $\hat{\mathcal{P}}_f \hat{\mathcal{P}}_p = 0$.

Los resultados previos muestran que $\hat{\mathcal{P}}_p$ y $\hat{\mathcal{P}}_f$ son proyectores que mandan a un campo al subespacio de los campos promedio y al subespacio de los campos fluctuantes respectivamente, y que el espacio donde vive originalmente el campo vectorial es una suma directa de estos dos subespacios. Esto permite escribir formalmente a un campo arbitrario como si fuera un vector de dos componentes,

$$\mathbf{F} = \begin{pmatrix} \mathbf{F}_p \\ \mathbf{F}_f \end{pmatrix}, \quad (9)$$

aunque cada componente en sí no es un número sino un campo vectorial. Análogamente, las funciones respuesta pueden representarse como operadores lineales en términos de matrices de dos por dos. Así, la ecuación $\mathbf{D} = \hat{\epsilon} \mathbf{E}$ puede escribirse como

una ecuación matricial

$$\begin{pmatrix} \mathbf{D}_p \\ \mathbf{D}_f \end{pmatrix} = \begin{pmatrix} \hat{\epsilon}_{pp} & \hat{\epsilon}_{pf} \\ \hat{\epsilon}_{fp} & \hat{\epsilon}_{ff} \end{pmatrix} \begin{pmatrix} \mathbf{E}_p \\ \mathbf{E}_f \end{pmatrix}. \quad (10)$$

Aquí, hemos definido $\hat{\mathbf{O}}_{\alpha\beta} = \hat{\mathcal{P}}_\alpha \hat{\mathbf{O}} \hat{\mathcal{P}}_\beta$ ($\alpha, \beta = p, f$) para cualquier operador $\hat{\mathbf{O}}$. Interpretaremos a la ec. (10) como un ecuación material *microscópica*, pues incorpora las fluctuaciones espaciales derivadas de la textura del material. La correspondiente ecuación macroscópica sería

$$\mathbf{D}_M = \hat{\epsilon}_M \mathbf{E}_M, \quad (11)$$

donde identificamos a los campos macroscópicos como los campos promediados y por lo tanto, libres de fluctuaciones, $\mathbf{D}_M \equiv \mathbf{D}_p$, $\mathbf{E}_M \equiv \mathbf{E}_p$. En general, $\hat{\epsilon}_M$ no es el promedio $\hat{\epsilon}_{pp}$ de $\hat{\epsilon}$, pues puede haber *correlaciones* entre las fluctuaciones espaciales de $\epsilon(\mathbf{r})$ y del campo eléctrico $\mathbf{E}(\mathbf{r})$.

2.2. Projectores longitudinal y transversal

Por otro lado, recordemos que de acuerdo al teorema de Helmholtz, todo campo vectorial $\mathbf{F} = \mathbf{F}^L + \mathbf{F}^T$ puede escribirse como la suma de dos contribuciones, un campo longitudinal $\mathbf{F}^L = \hat{\mathcal{P}}^L \mathbf{F}$ y un campo transversal $\mathbf{F}^T = \hat{\mathcal{P}}^T \mathbf{F}$, que cumplen las ecuaciones

$$\nabla \times \mathbf{F}^L = 0, \quad (12a)$$

$$\nabla \cdot \mathbf{F}^L = \nabla \cdot \mathbf{F}, \quad (12b)$$

$$\nabla \times \mathbf{F}^T = \nabla \times \mathbf{F}, \quad (12c)$$

$$\nabla \cdot \mathbf{F}^T = 0. \quad (12d)$$

Para obtener los proyectores $\hat{\mathcal{P}}^L$ y $\hat{\mathcal{P}}^T$ podemos empezar con la ec. (12a), la cual implica que \mathbf{F}^L puede derivarse de algún potencial escalar Φ como $\mathbf{F}^L = -\nabla\Phi$. Luego, la ecuación (12b) implica que el potencial obedece la ec. de Poisson, $\nabla^2\Phi = -\nabla \cdot \mathbf{F}$, cuya solución formal es $\Phi = -\hat{\nabla}^{-2} \hat{\nabla} \cdot \mathbf{F}$. Aquí, $\hat{\nabla}^{-2}$ representa el operador inverso al laplaciano, el cual puede escribirse en el espacio real como un operador integral cuyo kernel en 3D es el potencial coulombiano $-1/(4\pi|\mathbf{r} - \mathbf{r}'|)$. Finalmente, obtenemos $\mathbf{F}^L = \hat{\nabla} \hat{\nabla}^{-2} \hat{\nabla} \cdot \mathbf{F}$ de donde identificamos al proyector longitudinal

$$\hat{\mathcal{P}}^L = \hat{\nabla} \hat{\nabla}^{-2} \hat{\nabla} \cdot. \quad (13)$$

Análogamente, podemos identificar al proyector transversal

$$\hat{\mathcal{P}}^T = \hat{\mathbf{1}} - \hat{\mathcal{P}}^L = -\hat{\nabla} \times \hat{\nabla}^{-2} \hat{\nabla} \times. \quad (14)$$

2.3. Caso no retardado

Para obtener la respuesta macroscópica $\hat{\epsilon}_M$, que como hemos mencionado no es el simple promedio de la respuesta microscópica $\hat{\epsilon}$, recurrimos a las ecuaciones de Maxwell. Consideremos un material formado por inclusiones muy pequeñas y muy cercanas entre sí, cuyo tamaño y separación sean mucho menores que la longitud de onda característica de la luz $\lambda_0 = 2\pi c/\omega$ a una frecuencia ω dada. En este

caso podemos ignorar el *retardamiento*, tomar el límite $c \rightarrow \infty$ en la ecuación de Faraday y tratar al campo eléctrico como si fuera un campo puramente *longitudinal*, $\mathbf{E} = \mathbf{E}^L$, derivable de un potencial escalar. Por tanto, la proyección longitudinal del desplazamiento, $\mathbf{D}^L = \hat{\mathcal{P}}^L \hat{\epsilon} \mathbf{E} = \hat{\mathcal{P}}^L \hat{\epsilon} \mathbf{E}^L = \hat{\mathcal{P}}^L \hat{\epsilon} \hat{\mathcal{P}}^L \mathbf{E}^L$,

$$\mathbf{D}^L = \hat{\epsilon}^{LL} \mathbf{E}^L, \quad (15)$$

cumple las mismas ecuaciones que el campo eléctrico longitudinal externo,

$$\begin{aligned} \nabla \times \mathbf{D}^L &= 0 = \nabla \times \mathbf{E}_{\text{ex}}^L, \\ \nabla \cdot \mathbf{D}^L &= 4\pi \rho^{\text{ex}} = \nabla \cdot \mathbf{E}_{\text{ex}}^L. \end{aligned} \quad (16)$$

Aquí, hemos definido $\hat{\mathcal{O}}^{\alpha\beta} = \hat{\mathcal{P}}^\alpha \hat{\mathcal{O}} \hat{\mathcal{P}}^\beta$ ($\alpha, \beta = L, T$) para cualquier operador $\hat{\mathcal{O}}$. Entonces, podemos identificar a \mathbf{D}^L con el campo eléctrico longitudinal externo. Siendo un campo externo, sus fuentes son únicamente las cargas externas ρ^{ex} , las cuales no tienen absolutamente nada que ver con la composición del material ni con la disposición de sus componentes. En particular, \mathbf{D}^L no tiene fluctuaciones espaciales derivadas de la textura del material, $\mathbf{D}^L = \mathbf{D}_p^L = \hat{\mathcal{P}}_p \mathbf{D}^L$. Despejando el campo eléctrico de la ec. (15) obtenemos

$$\mathbf{E}^L = (\hat{\epsilon}^{LL})^{-1} \mathbf{D}^L, \quad (17)$$

y promediando ambos lados de la ecuación, usando el hecho de que \mathbf{D}^L no tiene fluctuaciones y que $\hat{\mathcal{P}}_p$ es idempotente, obtenemos

$$\mathbf{E}_p^L = (\hat{\epsilon}^{LL})_{pp}^{-1} \mathbf{D}_p^L. \quad (18)$$

Finalmente, identificando los campos promedios con los campos macroscópicos y su relación con la respuesta macroscópica, podemos identificar[3]

$$(\hat{\epsilon}_M^{LL})^{-1} = (\hat{\epsilon}^{LL})_{pp}^{-1}. \quad (19)$$

Podemos leer este resultado de la siguiente manera: el inverso de la proyección longitudinal de la respuesta dieléctrica macroscópica es igual al promedio del inverso de la proyección longitudinal de la respuesta microscópica.[7]

2.4. Caso retardado

Como hemos mostrado, para obtener la respuesta macroscópica de un sistema no basta con promediar cualquier función respuesta. Por ejemplo, el promedio de ϵ no tiene significado como respuesta macroscópica. Sin embargo, si logramos encontrar un operador que responda a una excitación *externa*, la cual no tiene fluctuaciones espaciales asociadas a la textura microscópica del material, su promedio nos proporciona la respuesta correspondiente macroscópica. Para dar un ejemplo de éste proceso, a continuación obtendremos la respuesta macroscópica en el caso en que no podemos ignorar el retardamiento. Tomando el rotacional de la ley de Faraday y sustituyendo la ec. de Ampère-Maxwell podemos obtener una ecuación de onda con fuentes, que podemos escribir como

$$\hat{\mathcal{W}} \mathbf{E} = \frac{4\pi}{i\omega} \mathbf{j}^{\text{ex}}, \quad (20)$$

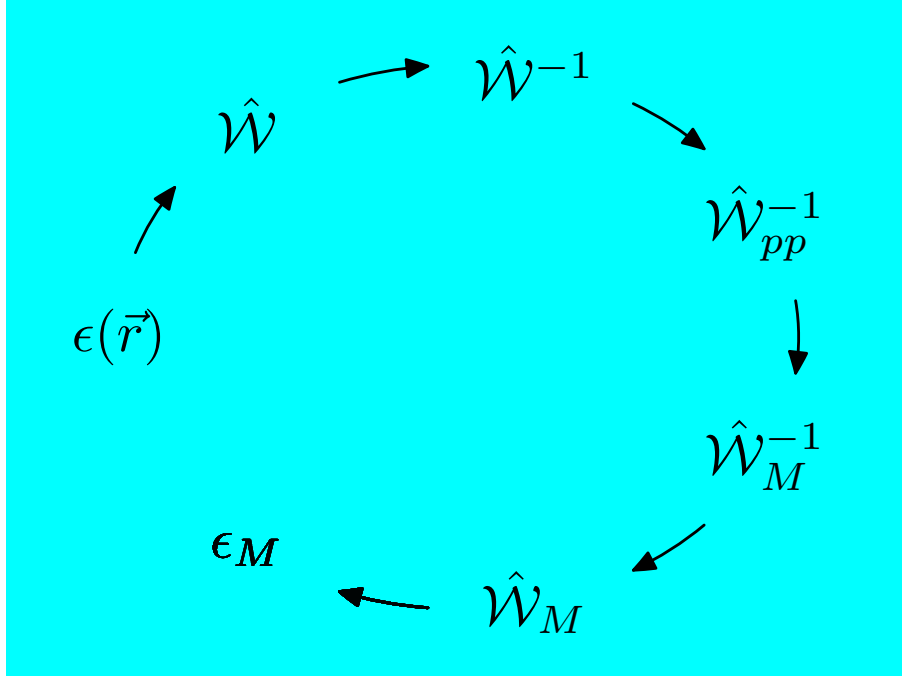


Figura 14: Pasos para obtener la respuesta macroscópica $\hat{\epsilon}_M$ a partir de la respuesta microscópica $\epsilon(\mathbf{r})$ incluyendo efectos de retardamiento. Primero construimos el operador de onda, lo invertimos, lo promediamos, identificamos el operador de onda macroscópico inverso, lo invertimos y finalmente identificamos el operador dieléctrico macroscópico.

donde

$$\hat{\mathcal{W}} = \hat{\epsilon} + \frac{c^2}{\omega^2} \nabla^2 \hat{\mathcal{P}}^T \quad (21)$$

es una generalización del *operador de onda*. Podemos resolver la ec. (20) formalmente para obtener el campo en el material

$$\mathbf{E} = \frac{4\pi}{i\omega} \hat{\mathcal{W}}^{-1} \mathbf{j}^{\text{ex}}, \quad (22)$$

e interpretar al inverso del operador de onda $\hat{\mathcal{W}}^{-1}$ como una respuesta a la excitación externa \mathbf{j}^{ex} , y a su promedio como la respuesta macroscópica a la corriente externa,

$$\hat{\mathcal{W}}_M^{-1} = \hat{\mathcal{W}}_{pp}^{-1}. \quad (23)$$

Interpretando al operador de onda macroscópico en términos de la permitividad macroscópica

$$\hat{\mathcal{W}}_M = \hat{\epsilon}_M + \frac{c^2}{\omega^2} \nabla^2 \hat{\mathcal{P}}^T \quad (24)$$

podemos despejarla tras invertir la ecuación (23). Formalmente, podemos obtener la permitividad macroscópica a partir de la permitividad microscópica siguiendo los pasos indicados en la figura 14. A partir de la respuesta dieléctrica de las componentes de nuestro material podemos construir el operador de onda, lo invertimos, lo promediamos, lo identificamos en términos del operador de onda macroscópico, lo volvemos a invertir y finalmente obtenemos la respuesta dieléctrica macroscópica.[8]

2.5. Sistema binario periódico sin retardamiento

Consideremos ahora un sistema periódico hecho de dos materiales, digamos, de partículas de un material B embebidas en una matriz de un material A . La respuesta dieléctrica *microscópica* es en este caso

$$\epsilon(\mathbf{r}) = \begin{cases} \epsilon_A & \text{si } \mathbf{r} \in A \\ \epsilon_B & \text{si } \mathbf{r} \in B \end{cases} \quad (25)$$

la cual puede escribirse como

$$\epsilon(\mathbf{r}) = \frac{\epsilon_A}{u}(u - B(\mathbf{r})), \quad (26)$$

donde

$$u \equiv \frac{1}{1 - \epsilon_B/\epsilon_A} \quad (27)$$

se conoce como la *variable espectral*, y

$$B(\mathbf{r}) = \begin{cases} 0, & \text{si } \mathbf{r} \in A \\ 1, & \text{si } \mathbf{r} \in B \end{cases} \quad (28)$$

es la *función característica*. Notemos que u depende de la composición del material y puede depender de la frecuencia a través de las funciones dieléctricas de las componentes, mientras que $B(\mathbf{r})$ no depende de la composición ni de la frecuencia, sino únicamente de la geometría.

Supongamos que nuestro sistema es periódico, caracterizado por una *red de Bravais* $\{\mathbf{R}\}$, tal que $\epsilon(\mathbf{r} + \mathbf{R}) = \epsilon(\mathbf{r})$. A esta red le corresponde una *red recíproca* $\{\mathbf{K}\}$, formada por todos aquellos vectores de onda tales que el producto escalar $\mathbf{K} \cdot \mathbf{R} = 2\pi n$, con n un número entero para cualquier vector real \mathbf{R} y vector recíproco \mathbf{K} .

Entonces, podemos describir a $\hat{\epsilon}$ como una matriz en el espacio recíproco mediante una integral de Fourier,

$$\epsilon_{\mathbf{K}\mathbf{K}'} = \epsilon_{\mathbf{K}-\mathbf{K}'} = \frac{1}{\Omega} \int_0 d^3r \epsilon(\mathbf{r}) e^{-i\mathbf{K}\cdot\mathbf{r}}. \quad (29)$$

donde la integral se realiza sobre una celda primitiva cualquiera, cuyo volumen es Ω . La relación conversas es

$$\epsilon(\mathbf{r}) = \sum_{\mathbf{K}} \epsilon_{\mathbf{K}} e^{i\mathbf{K}\cdot\mathbf{r}}. \quad (30)$$

El teorema de Bloch nos permite escribir a los campos en el interior de un sistema periódico como una superposición de ondas de Bloch, cada una de las cuales cambia por una fase $e^{i\mathbf{k}\cdot\mathbf{R}}$ cuando nos desplazamos un vector de la red, donde \mathbf{k} es el *vector de Bloch*, una especie de vector de onda, que es una *cantidad conservada*. Por tanto, eligiendo un valor cualquiera de \mathbf{k} podemos escribir cualquier campo en términos de una suma de Fourier como

$$\mathbf{F}(\mathbf{r}) = \sum_{\mathbf{K}} \mathbf{F}_{\mathbf{K}} e^{i(\mathbf{k}+\mathbf{K})\cdot\mathbf{r}}. \quad (31)$$

El *teorema de convolución* nos permite entonces escribir la ecuación material $\mathbf{D}(\mathbf{r}) = \epsilon(\mathbf{r})\mathbf{E}(\mathbf{r})$ en el espacio recíproco como la ecuación matricial

$$\mathbf{D}_{\mathbf{K}} = \sum_{\mathbf{K}'} \epsilon_{\mathbf{K}\mathbf{K}'} \mathbf{E}_{\mathbf{K}'}. \quad (32)$$

Como la respuesta del material puede representarse en distintos espacios, como son el espacio real o el espacio recíproco, conviene abstraer la acción de la respuesta dieléctrica y tratarla como un operador $\hat{\epsilon}$ abstracto, más que como una función de la posición.

En el espacio recíproco, el operador ∇ se puede representar por un producto con el vector $\nabla \rightarrow i(\mathbf{k} + \mathbf{K})$, por lo cual el laplaciano es $\nabla^2 \rightarrow -|\mathbf{k} + \mathbf{K}|^2$ y su inverso es simplemente $\nabla^{-2} \rightarrow -1/|\mathbf{k} + \mathbf{K}|^2$. Por lo tanto, el proyector longitudinal (13) se puede representar por la matriz

$$\mathcal{P}_{\mathbf{K}\mathbf{K}'}^L = \delta_{\mathbf{K}\mathbf{K}'} \hat{\mathbf{K}} \hat{\mathbf{K}}', \quad (33)$$

donde definimos los vectores unitarios

$$\hat{\mathbf{K}} \equiv \frac{\mathbf{k} + \mathbf{K}}{|\mathbf{k} + \mathbf{K}|}, \quad (34)$$

y $\delta_{\mathbf{K}\mathbf{K}'}$ es una función delta de Kronecker.

Los vectores recíprocos $\mathbf{K} \neq 0$ corresponden a oscilaciones con longitudes de onda del orden de el parámetro de red del sistema periódico. Es conveniente entonces *definir* el promedio como un filtro pasabajos en el espacio recíproco que elimina todos los vectores recíprocos, exceptuando $\mathbf{K} = 0$, convirtiendo una onda de Bloch en una onda plana con vector de onda \mathbf{k} . Escribimos entonces

$$\mathcal{P}_{p,\mathbf{K}\mathbf{K}'} = \delta_{\mathbf{K}\mathbf{0}} \delta_{\mathbf{K}'\mathbf{0}}. \quad (35)$$

De acuerdo a la ec. (19), podemos hallar la respuesta macroscópica siguiendo los siguientes pasos:

1. Expresamos la permitividad como una matriz $\epsilon_{\mathbf{K}\mathbf{K}'}$ en el espacio recíproco haciendo una transformación de Fourier (ec. (29)).
2. Tomamos su proyección longitudinal, $\epsilon_{\mathbf{K}\mathbf{k}}^{LL}$ multiplicando a izquierda y derecha por el proyector longitudinal (ec. (33)).
3. Invertimos la matriz resultante en el subespacio de campos vectoriales longitudinales $(\epsilon_{\mathbf{K}\mathbf{K}'}^{LL})^{-1}$.
4. Promediamos la inversa tomando el elemento $\mathbf{K} = \mathbf{K}' = 0$.
5. Interpretamos el resultado como el inverso de la proyección longitudinal del tensor dieléctrico macroscópico.

El procedimiento anterior puede resumirse como $\epsilon_M^{LL} = \hat{\mathbf{k}} \epsilon_M^{LL} \hat{\mathbf{k}}$, donde la *componente longitudinal* de la respuesta es

$$\epsilon_M^{LL} = \hat{\mathbf{k}} \cdot \epsilon_M \cdot \hat{\mathbf{k}} = \left(\left(\hat{\mathbf{K}} \cdot \epsilon_{\mathbf{K}\mathbf{K}'} \hat{\mathbf{K}}' \right)^{-1} \Big|_{\mathbf{K}=\mathbf{K}'=0} \right)^{-1}. \quad (36)$$

Repitiendo el cálculo indicado por la ec. (36) para distintas direcciones $\hat{\mathbf{k}}$ del vector de onda, podemos hallar todas las componentes del tensor dieléctrico macroscópico ϵ_M .

2.6. Recursión de Haydock

De acuerdo a la sección anterior, para obtener la respuesta macroscópica de un medio binario arbitrario en el límite de longitud de onda larga, basta invertir la matriz $\hat{\mathbf{K}} \cdot \epsilon_{\mathbf{K}\mathbf{K}'} \hat{\mathbf{K}}'$ y tomar el elemento $\mathbf{00}$ del resultado. Sustituyendo la ec. (26), obtenemos

$$\frac{1}{\epsilon_M^{LL}} = \frac{u}{\epsilon_A} (u\delta_{\mathbf{K}\mathbf{K}'} - B_{\mathbf{K}\mathbf{K}'}^{LL})^{-1} \Big|_{\mathbf{K}=\mathbf{K}'=0}, \quad (37)$$

donde introdujimos la componente longitudinal de la función característica en el espacio recíproco

$$B_{\mathbf{K}\mathbf{K}'}^{LL} = \hat{\mathbf{K}} \cdot B_{\mathbf{K}-\mathbf{K}'} \hat{\mathbf{K}}', \quad (38)$$

y $B_{\mathbf{K}-\mathbf{K}'}$ es el coeficiente de Fourier de la función característica $B(\mathbf{r})$ correspondiente al vector recíproco $\mathbf{K} - \mathbf{K}'$.

Notamos que en la ec. (37) debemos calcular el inverso de un operador $(u - \hat{B}^{LL})$ representado como una matriz en el espacio recíproco, donde u es un número complejo y \hat{B}^{LL} es un operador hermitiano, y luego proyectar el resultado sobre un estado $|0\rangle$ correspondiente a una onda plana con vector de onda \mathbf{k} . Esto es análogo al cálculo del operador de Green proyectado $\langle 0|\hat{\mathcal{G}}(\varepsilon)|0\rangle$ en mecánica cuántica, donde $\hat{\mathcal{G}}(\varepsilon) = (\varepsilon - \hat{\mathcal{H}})^{-1}$ es el operador de Green correspondiente a un operador hamiltoniano \mathcal{H} evaluado para un valor ε de una energía compleja. Entre las aplicaciones del operador de Green proyectado se encuentra el cálculo de la densidad de estados cuánticos proyectada. Podemos entonces tomar prestado el método recursivo de Haydock [9] para calcular proyecciones de funciones de Green. De acuerdo a nuestra analogía, u juega el papel de energía compleja ε y \hat{B}^{LL} juega el papel de hamiltoniano $\hat{\mathcal{H}}$.

Hemos definido el estado $|0\rangle$ como el correspondiente a una onda plana con vector de onda \mathbf{k} , i.e., un estado cuya representación en el espacio recíproco es $\langle \mathbf{K}|0\rangle = \delta_{\mathbf{K}\mathbf{0}}$, pues no tiene contribuciones de ondas con $\mathbf{K} \neq 0$. Ahora, podemos generar un nuevo estado haciendo actuar a nuestro hamiltoniano sobre el estado inicial, $|\tilde{1}\rangle = \hat{\mathcal{H}}|0\rangle$. Escribimos este estado como una combinación lineal del estado que ya teníamos $|0\rangle$ y un estado nuevo $|1\rangle$ del cual pedimos que sea ortogonal a $|0\rangle$ y que esté normalizado, $|\tilde{1}\rangle = b_1|1\rangle + a_0|0\rangle$. Ahora repetimos el procedimiento con el estado $|1\rangle$, i.e., $|\tilde{2}\rangle = \hat{\mathcal{H}}|1\rangle = b_2|2\rangle + a_1|1\rangle + b_1|0\rangle$. El caso genérico sería

$$|\tilde{n}\rangle = \hat{\mathcal{H}}|n-1\rangle = b_n|n\rangle + a_{n-1}|n-1\rangle + b_{n-1}|n-2\rangle, \quad (39)$$

donde exigimos que todos los estados sean ortonormales, es decir,

$$\langle n|m\rangle = \delta_{nm}. \quad (40)$$

Aquí empleamos el producto escalar

$$\langle n|m\rangle = \sum_{\mathbf{K}} \phi_n^*(\mathbf{K})\phi_m(\mathbf{K}) = \frac{1}{\Omega} \int_0 d^3r \phi_n^*(\mathbf{r})\phi_m(\mathbf{r}) \quad (41)$$

donde, copiando el lenguaje de la mecánica cuántica, $\phi_m(\mathbf{K}) = \langle \mathbf{K}|m\rangle$ es la *función de onda* correspondiente al estado $|n\rangle$ evaluada en el vector recíproco \mathbf{K} , y $\phi_m(\mathbf{r}) =$

$\langle \mathbf{r}|n\rangle$ es la función de onda correspondiente al mismo estado pero evaluada en la posición \mathbf{r} , y donde hemos empleado la identidad de Parseval.

Notamos que en la ec. (39) no aparecen los términos $|n-3\rangle$, $|n-4\rangle$, etc. pues nuestro operador es hermitiano. Por ejemplo, i.e., $\langle n-3|\tilde{n}\rangle = \langle n-3|H|n-1\rangle = \langle n-1|H|n-3\rangle^* = (\langle n-1|(b_{n-2}|n-2\rangle + a_{n-3}|n-3\rangle + b_{n-3}|n-4\rangle))^* = 0$. Los coeficientes de Haydock son reales y pueden obtenerse de la condición de ortonormalidad,

$$a_{n-1} = \langle n-1|\hat{\mathcal{H}}|n-1\rangle \quad (42)$$

y

$$|n\rangle = (|\tilde{n}\rangle - a_{n-1}|n-1\rangle - b_{n-1}|n-1\rangle)/b_n, \quad \langle n|n\rangle = 1. \quad (43)$$

De esta manera podemos construir una *base* $\{|n\rangle\}$ en la cual el *hamiltoniano* $\hat{\mathcal{H}}$ puede representarse por una matriz tridiagonal

$$\mathcal{H}_{nn'} = \begin{pmatrix} a_0 & b_1 & 0 & 0 & 0 \dots \\ b_1 & a_1 & b_2 & 0 & 0 \dots \\ 0 & b_2 & a_2 & b_3 & 0 \dots \\ 0 & 0 & b_3 & a_3 & b_4 \dots \\ \vdots & \vdots & \vdots & \ddots & \ddots \ddots \end{pmatrix} \quad (44)$$

De acuerdo a la ec. (37), la respuesta macroscópica está dada por la proyección de la inversa $\mathcal{G}_{nn'}$ de la matriz tridiagonal $\varepsilon - \mathcal{H}_{nn'}$ sobre el estado $|0\rangle$, i.e., no necesitamos toda la inversa, sino sólo su elemento \mathcal{G}_{00} . Notamos que el vector columna \mathcal{G}_{n0} obedece la ecuación

$$\sum_m (\varepsilon \delta_{nm} - \mathcal{H}_{nm}) \mathcal{G}_{m0} = \delta_{n0}. \quad (45)$$

Si truncamos la ecuación después de N renglones, el último renglón de esta ecuación sería de la forma

$$-b_N \mathcal{G}_{N-1,0} + (u - a_N) \mathcal{G}_{N,0} = 0, \quad (46)$$

lo cual nos permite despejar

$$\mathcal{G}_{N,0} = \frac{b_N}{u - a_N} \mathcal{G}_{N-1,0}. \quad (47)$$

Sustituyendo esta solución en la penúltima ecuación

$$\begin{aligned} & -b_{N-1} \mathcal{G}_{N-2,0} + (u - a_{N-1}) \mathcal{G}_{N-1,0} - b_N \mathcal{G}_{N,0} \\ & = -b_{N-1} \mathcal{G}_{N-2,0} + \left(u - a_{N-1} - \frac{b_N^2}{u - a_N} \right) \mathcal{G}_{N-1,0} = 0, \end{aligned} \quad (48)$$

podemos despejar

$$\mathcal{G}_{N-1,0} = \frac{b_{N-1}}{u - a_{N-1} - \frac{b_N^2}{u - a_N}} \mathcal{G}_{N-2,0}. \quad (49)$$

Prosiguiendo de esta forma con todas las ecuaciones correspondientes a $n > 0$,

$$-b_n \mathcal{G}_{n-1,0} + (u - a_n) \mathcal{G}_{n,0} - b_{n+1} \mathcal{G}_{n+1,0} = 0, \quad (50)$$

llegamos a

$$\mathcal{G}_{1,0} = \frac{b_1}{u - a_1 - \frac{b_2^2}{u - a_2 - \frac{b_3^2}{u - a_3 - \ddots}}} \mathcal{G}_{0,0}. \quad (51)$$

Sustituyendo en la ecuación correspondiente a $n = 0$,

$$(u - a_0)\mathcal{G}_{00} - b_{n+1}\mathcal{G}_{1,0} = 1, \quad (52)$$

obtenemos una expresión para \mathcal{G}_{00} en forma de una fracción continuada

$$\mathcal{G}_{00} = \frac{1}{u - a_0 - \frac{b_1^2}{u - a_1 - \frac{b_2^2}{u - a_2 - \frac{b_3^2}{u - a_3 - \ddots}}}}, \quad (53)$$

la cual podemos emplear en la ec. (37) para finalmente obtener la respuesta macroscópica

$$\epsilon_M^{LL} = \frac{\epsilon_A}{u} \left(u - a_0 - \frac{b_1^2}{u - a_1 - \frac{b_2^2}{u - a_2 - \frac{b_3^2}{u - a_3 - \ddots}}} \right). \quad (54)$$

Debemos enfatizar que en esta expresión, los coeficientes de Haydock $\{a_n, b_n\}$ dependen exclusivamente de la geometría a través de la función característica $B(r)$ y no de la composición del material ni de la frecuencia. Por lo tanto, sólo es necesario calcularlos una vez para una geometría dada y posteriormente pueden emplearse para calcular la respuesta de cualquier metamaterial con dicha geometría con cualquier composición y a cualquier frecuencia, simplemente sustituyendo la variable espectral adecuada u (ec. (27)). Así, con este formalismo podemos calcular la respuesta de sistemas formados por aislantes o metales, con o sin dispersión y con o sin disipación; el operador \hat{B}^{LL} es hermitiano haya o no haya dispersión o disipación.

Por otro lado, podemos aplicar el hamiltoniano \hat{B}^{LL} (38) en etapas, notando que multiplicar por $\hat{\mathbf{K}}'$ es trivial en el espacio recíproco. El producto matricial del vector resultante con $B_{\mathbf{K}\mathbf{K}'} = B_{\mathbf{K}-\mathbf{K}'}$ corresponde a una convolución, por lo cual, tras una transformada de Fourier hacia el espacio real, se convierte en un producto trivial por la función característica $B(\mathbf{r})$. Finalmente, tomando una transformada de Fourier de regreso al espacio recíproco, el producto escalar por los vectores unitarios $\hat{\mathbf{K}}$ se vuelve trivial. Esto muestra que podemos aplicar nuestro hamiltoniano repetidas veces para obtener los coeficientes de Haydock sin necesidad de multiplicar ninguna matriz. Eso vuelve muy eficiente el proceso aquí descrito.

3. Implementación

La teoría mostrada arriba ha sido implementada en un paquete computacional llamado *Photonic*, el cual ha sido colocado en el *dominio público* [10, 11]. El programa está escrito en el lenguaje *PERL*, el cual es muy expresivo y cuya sintaxis

hereda construcciones de lenguajes previos como *C*, *C++*, *awk*, *bash*, *sed*, etc. El lenguaje es versátil y flexible, y refleja la filosofía de su creador resumida en frases como *hay más de una manera de resolverlo y lo fácil debe ser fácil, lo difícil debe ser posible*. Sólo contiene tres tipos de datos, que son *escalares*, *arreglos* indexados por un entero y *arreglos asociativos* indexados por cualquier escalar. Los escalares pueden representar números enteros o reales, cadenas de caracteres, o referencias a otros escalares, arreglos o arreglos asociativos, e incluso, referencias a subrutinas. Los arreglos son dinámicos y pueden crecer o decrecer en ambos extremos o en su interior. Esto permite implementar de manera trivial pilas de datos (*stacks*), colas *fifo*, árboles y otras estructuras de datos. Se pueden construir fragmentos de código durante la ejecución de un programa para ejecutarse posteriormente. Esta flexibilidad permite emplear una gran variedad de paradigmas al programar en *PERL*, incluyendo programación procedural, funcional y/o orientada a objetos.

El costo a pagar por la flexibilidad del lenguaje es la velocidad de ejecución. Se ha reportado recientemente que *PERL* es varias decenas de veces más lento que *C* para tareas orientadas a procesamiento numérico.[12] Por ello, un grupo de investigadores se abocaron a crear una extensión del lenguaje llamado *Perl Data Language* o *PDL* [13, 14, 15], que permite ligar rutinas numéricas escritas en otros lenguajes como *C* o *Fortran* para el manejo eficiente de arreglos numéricos, sin sacrificar la flexibilidad y expresividad de *PERL*. Una prueba reciente[12] mostró que *PDL* es competitivo y hasta puede superar en velocidad a códigos nativos en *C*.

Finalmente, para simplificar el proceso de codificación y volverlo robusto conforme evoluciona el paquete se empleó un sistema de programación de objetos conocido como *Moose*[16]. Este sistema permite definir *clases* que abstraen el comportamiento de los *objetos*, instancias que tienen una serie de *atributos*, datos privados, y *métodos* que definen su comportamiento. Las clases pueden heredar su comportamiento de otras clases o de *roles* que definen las interfases. Contar con una librería de clases permite armar programas que resuelven problemas complejos juntando bloques que ensamblan unos con otros, como las construcciones con bloques de juguetes *Lego*.

3.1. Ejemplo

No explicaremos aquí los detalles de la implementación, pues su comprensión requeriría cierto dominio de los sistemas (Perl, PDL y Moose) empleados. En cambio, mostraremos fragmentos de un pequeño programa para explicar cómo se usa el sistema. El programa íntegro, disponible en la referencia [17], calcula el tensor dieléctrico de una red tetragonal de toroides hechos de cierto material y embebido en una matriz de otro material.

Iniciamos con una serie de *pragmas* y cargando paquetes que serán útiles más adelante

```
#!/usr/bin/env perl
# ...
use strict;
use warnings;
use v5.12;
use Getopt::Long;
```

```

use PDL;
use PDL::NiceSlice;
use PDL::Constants qw(PI);
use Photonic::Geometry::FromB;
use Photonic::LE::NR2::Haydock;
use Photonic::LE::NR2::EpsL;

```

`strict` y `warnings` son para pedir al sistema que sea estricto y nos advierta de errores potenciales, `v5.12` es para habilitar algunas construcciones semánticas, `Getopt` es para leer los parámetros desde la línea de comandos al ejecutar el programa, `PDL` es para usar la interface numérica, `NiceSlice` para simplificar el manejo de índices en las estructuras de datos y `PI` es simplemente una constante útil. Los paquetes relacionados con `Photonic` serán discutidos más abajo. Las componentes `LE` y `NR2` en su nombre indican que usaremos aquellas rutinas relacionadas con la respuesta dieléctrica longitudinal en el límite no retardado y restringido a dos componentes.

A continuación definimos algunos parámetros y el código para leerlos desde la línea de comandos.

```

my $ratio; # b/a for torus
...
my $options=q(
'ratio=f'=>\$ratio,
...
);
...
GetOptions( %options)or usage($options, "Bad options");
usage($options, "Missing options")
    unless !all {defined $_}
        ($ratio, $fraction, @eps_a, @eps_b, $Nxy, $Nz, $Nh);
...
set_autopthread_targ($cores) if defined $cores;;

```

Los parámetros a leer son la razón de los radios del toroide, la fracción de llenado en la celda unitaria, el numero de *voxels* a lo largo de los ejes de la red, los pares de funciones dieléctricas e emplear para los toroides y la matriz, el número de coeficientes de Haydock a emplear y el número de núcleos computacionales a emplear en el cálculo. A continuación se leen y validan las opciones y de ser necesario se envían mensajes de error. La rutina `set_autopthread_targ` establece el número de núcleos computacionales que deseamos usar al paralelizar el programa.

Como indicamos arriba, la geometría queda definida a partir de la función característica, cuyo valor es 1 dentro del toroide y 0 en su exterior. Primero calculamos los dos radios del toroide en términos de la fracción de llenado deseada.

```

my ($Nxy2, $Nz2)=(2*$Nxy+1, 2*$Nz+1);
my $unit_cell_volume=$Nxy2*$Nxy2*$Nz2;
my $small_radius=($fraction*$unit_cell_volume/(2*PI**2*$ratio))**(1/3);
my $large_radius=$ratio*$small_radius;
warn "Tori overlap" if $small_radius>$Nz
                    or $large_radius+$small_radius>$Nxy;

```

Luego creamos un arreglo 3D representando a la celda unitaria y lo poblamos de unos y ceros de acuerdo a la función característica deseada.

```
my $r=zeros($Nxy2, $Nxy2, $Nz2)->ndcoords
    -pdl($Nxy, $Nxy, $Nz); #positions array
my $B=(sqrt($r((0))**2+$r((1))**2)-$large_radius)**2
    +$r((2))**2 < $small_radius**2;
```

La rutina `zeros` produce el arreglo 3D de ceros con el número de dimensiones y el tamaño solicitado, el *método* `ndcoords` asigna a cada punto del arreglo un vector en 3D con las coordenadas de dicho punto. Al restar las coordenadas del centro del arreglo, asignamos a la variable `$r` las coordenadas de los puntos del arreglo con respecto al centro. Procesando dichas coordenadas obtenemos la distancia de cada punto al círculo alrededor del cual se forma el toroide y en la variable `$B` guardamos un 1 o un 0 dependiendo de si estamos suficientemente cerca de la generatriz o no.

A continuación inicializamos dos objetos que codifican la geometría del sistema empleando la función característica.

```
my $gx=Photonic::Geometry::FromB->new(B=>$B, Direction0=>pdl(1,0,0));
my $gz=Photonic::Geometry::FromB->new(B=>$B, Direction0=>pdl(0,0,1));
```

Las clases `Geometry` saben cómo calcular la red de *voxels*, pero también cómo calcular la red recíproca, y los vectores recíprocos normalizados \hat{K} , entre otros métodos relacionados a la geometría del sistema, a partir del atributo `B` que inicializamos con la función característica `$B`. Sin embargo, para ello necesita saber en qué dirección apunta el vector de Bloch, asociado al atributo `Direction0`. Emplearemos entonces dos *geometrías*, una con ondas viajando en la dirección \hat{x} y otra viajando en la dirección \hat{z} , i.e., a lo largo del plano y del eje del toroide, respectivamente.

Con las dos geometrías podemos inicializar dos objetos para calcular coeficientes de Haydock.

```
my $nrx=Photonic::LE::NR2::Haydock->new(geometry=>$gx, nh=>$Nh);
my $nrz=Photonic::LE::NR2::Haydock->new(geometry=>$gz, nh=>$Nh);
```

Finalmente, para cada pareja de funciones dieléctricas, calculamos la respuesta dieléctrica macroscópica longitudinal, proyectada sobre la dirección del vector de Bloch establecida arriba.

```
say "#ratio Nxy Nz Nh f-nom f-act medium torus epsxx epszz";
foreach(0..@eps_a-1){
    my ($ea, $eb)=(pdl($eps_a[$_])->r2C, pdl($eps_b[$_])->r2C);
    my $epsx_calc=Photonic::LE::NR2::EpsL->new(haydock=>$nrx, nh=>$Nh,
        epsA=>$ea, epsB=>$eb);
    my $epsz_calc=Photonic::LE::NR2::EpsL->new(haydock=>$nrz, nh=>$Nh,
        epsA=>$ea, epsB=>$eb);
    my $resultx=$epsx_calc->epsL;
    my $resultz=$epsz_calc->epsL;
```

El objeto `EpsL` sabe calcular la respuesta dieléctrica longitudinal y se inicializa con un objeto que calcula coeficientes de Haydock, con el número de coeficientes que se desea usar y con las funciones dieléctricas de ambas componentes. Finalmente, el método `epsL` regresa el *valor* de la función dieléctrica macroscópica deseada.

El resto del programa simplemente imprime el resultado y mensajes de error de ser necesario.

```

say sprintf "%.4f %d %d %d %.4f %.4f %.4f %.4f %.4f %.4f",
$ratio, $Nxy, $Nz, $Nh, $fraction, $gx->f, $ea->re, $eb->re,
$resultx->re, $resultz->re;
say "x-no-covergió" unless $epsx_calc->converged;
say "z-no-covergió" unless $epsz_calc->converged;
}

sub usage {
    ...
}

```

Podemos correr el programa como en el siguiente ejemplo, en el que calculamos las propiedades de una red tetragonal de $161 \times 161 \times 41$ voxeles con toroides cuyas funciones dieléctricas son $\epsilon_b = 5, 10$, inmersos en el vacío, $\epsilon_a = 1$, con una fracción de llenado nominal $f = 0.3$, con una razón entre los radios del círculo mayor al círculo menor $R_>/R_< = 3$ y empleando 100 coeficientes de Haydock, usando en el cálculo 4 núcleos de la unidad de procesamiento,

```

./toroid.pl -ratio 3 -fraction .3 -Nz 20 -Nxy 80 \
    -eps_a 1 -eps_b 5 -eps_a 1 -eps_b 10 -Nh 100 -cores 4

```

obteniendo unos segundos después la siguiente tabla:

#ratio	Nxy	Nz	Nh	f-nom	f-act	medium	torus	epsxx	epszz
3.0000	80	20	100	0.3000	0.3004	1.0000	5.0000	1.7228	1.6859
3.0000	80	20	100	0.3000	0.3004	1.0000	10.0000	2.1836	2.0152

Las últimas dos columnas nos proporcionan las componentes del tensor dieléctrico macroscópico de este sistema.

3.2. Extensiones

Además de la teoría no retardada desarrollada en detalle en la sección 2.5, hemos desarrollado la teoría para poder calcular la respuesta dieléctrica en presencia de retardamiento (sección 2.4) y para sistemas no binarios, con tres o más componentes. También hemos extendido la teoría para poder calcular los campos electromagnéticos microscópicos y a partir de ellos calcular propiedades no lineales. Estas extensiones han sido incorporadas en el paquete *Photonic*.

3.3. Instalación

Para instalar el paquete *Photonic* es necesario instalar primero el paquete *PDL*. En un sistema *linux* basta emplear el comando `cpanm -look PDL` desde una línea de comandos. Debe leerse entonces el archivo `INSTALLATION` e instalar su lista de prerequisites manualmente para posteriormente invocar los comandos `perl Makefile.PL`, `make`, `make test` y `make install`. Posteriormente, el comando `cpanm Photonic` instala de manera automática nuestro sistema. Una vez instalado, el comando `perldoc Photonic` da acceso al manual en línea.

4. Resultados

En esta sección enumeraremos algunos de los resultados que hemos obtenido con la teoría y códigos descritos arriba.

4.1. Dicroísmo lineal y transmisión extraordinaria

El punto de inicio de nuestros cálculos es la función característica $B(\mathbf{r})$, consistente en unos y ceros, dependiendo de si \mathbf{r} se halla dentro del material B o A . Al discretizar el espacio real en 2D, $B(\mathbf{r})$ se vuelve una representación binaria de una imagen pixelada. Por lo tanto, nuestro programa puede alimentarse *literalmente* de una imagen pixelada de alto contraste. Esto permite manipular la imagen usando herramientas gráficas y calcular las propiedades ópticas del sistema resultante. Ilustramos esto con la fig. 15, en cuyo lado izquierdo mostramos un corte de una red rectangular de agujeros cilíndricos con sección transversal elíptica en una matriz de plata. Manipulamos gráficamente la razón de aspecto de la red, y la excentricidad y orientación de las elipses. En la figura hemos elegido el valor 2 para la razón de aspecto de la red y escogimos 1.8 para la razón entre los semiejes de las elipses. En el lado derecho mostramos el espectro de reflectancia a incidencia normal de una película delgada, de 100\AA de ancho, formada por este metamaterial. Mostramos dos conjuntos de datos, pues el material es anisótropo. Para cierta polarización, casi horizontal, el material se comporta como un metal ordinario con una alta reflectancia, cercana a $R = 1$. Sin embargo, para una polarización ortogonal, casi vertical hay un espectro de reflectancia con un mínimo profundo alrededor de $\hbar\omega = 2.4eV$, en que nuestra película muestra un dicroísmo extremo, i.e., casi toda la luz se refleja para una polarización y casi nada para la polarización ortogonal.[18] El mínimo de reflectancia puede entonarse a través de toda la región visible mediante cambios pequeños en la orientación θ de las elipses. El motivo del dicroísmo extremo es que en la dirección vertical los caminos conductores se hallan casi estrangulados, pero abiertos. Por tanto, mientras que el sistema es un buen conductor para campos horizontales, es un mal conductor para campos verticales y para campos que oscilan rápidamente el sistema se comporta como un dieléctrico con resonancias relacionadas a los plasmones localizados en los cilindros. Por tanto, la permitividad en la dirección vertical es negativa a bajas frecuencias (comportamiento metálico) pero positiva y con resonancias a altas frecuencias (comportamiento dieléctrico). Interpolando entre ambos comportamientos, para alguna frecuencia intermedia la respuesta empata con la del

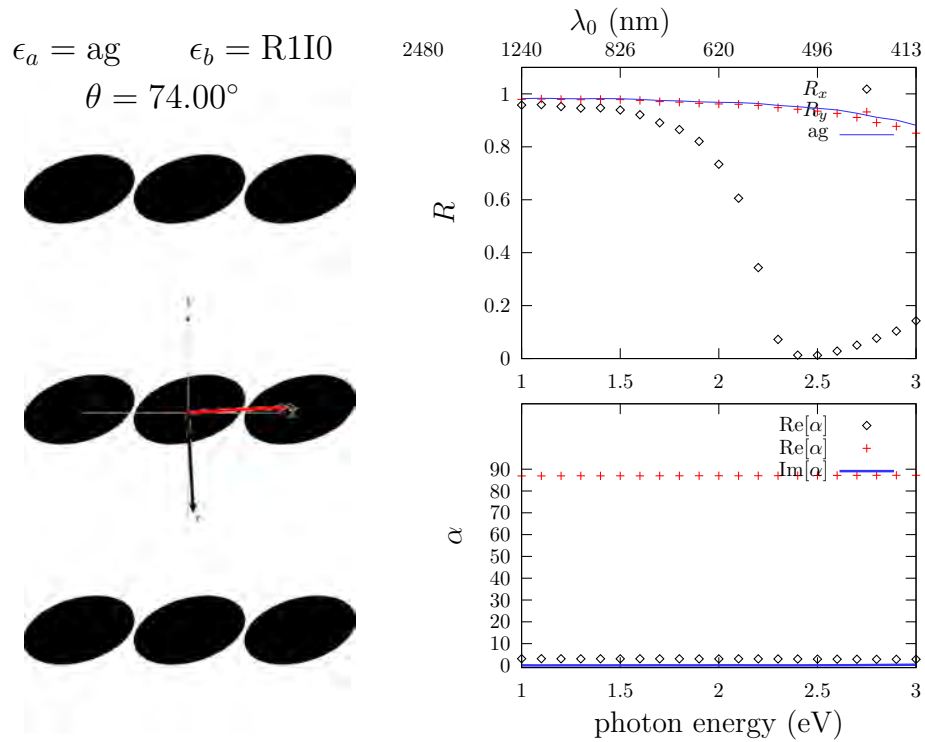


Figura 15: (Izquierda) Red rectangular con parámetros de red $L_y = 2L_x$ de agujeros cilíndricos con razón de semiejes $a = 1.8b$ en una matriz de plata. El semieje mayor está rotado $\theta = 74^\circ$ con respecto a la vertical. (Derecha arriba) Espectro de reflectancia de una película delgada de ancho 100\AA iluminada normalmente por luz linealmente polarizada a lo largo de los ejes principales del tensor dieléctrico, cuya dirección se indica a la izquierda y a la derecha abajo. Como referencia, se muestra la reflectancia de una película homogénea con la misma cantidad de Ag (línea continua). (Derecha abajo) Direcciones principales del tensor dieléctrico.

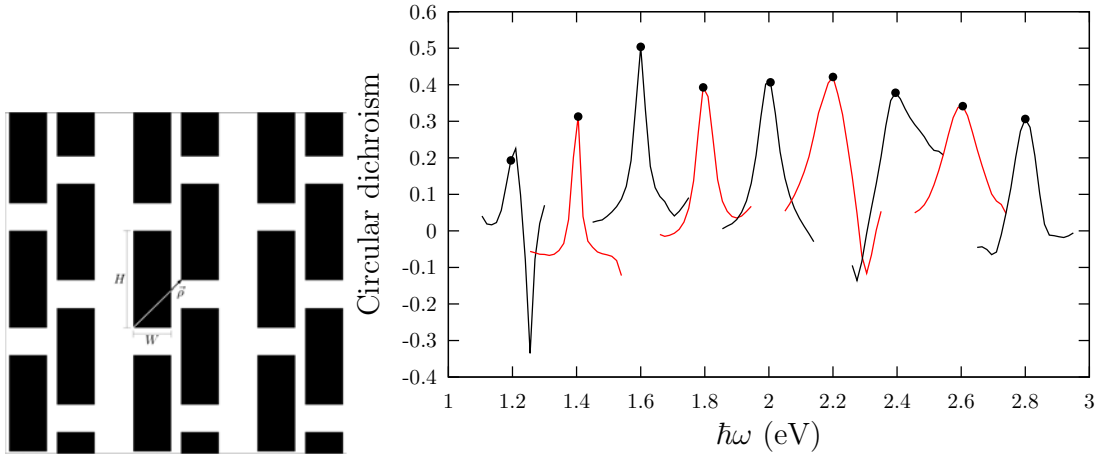


Figura 16: (Izquierda) Película de Ag depositada en vidrio con una red cuadrada de parejas de agujeros en forma de prismas rectangulares rellenos de un dieléctrico con permitividad ϵ_b . El sistema está parametrizado por el ancho W y altura H de los prismas y el desplazamiento ρ entre parejas de agujeros en la celda unitaria. (Derecha) Dicroísmo circular de películas de 100nm de ancho con parámetros entonados para maximizar el dicroísmo a una frecuencia preestablecida (puntos).

vacío y el material adquiere una transparencia extraordinaria, lo cual contrasta con la alta reflectancia para una polarización ortogonal.

4.2. Dicroísmo circular

En la fig. 16 mostramos un sistema formado por una red cuadrada de parejas de agujeros en forma de prismas rectangulares en una película delgada de Ag colocada sobre un sustrato de vidrio. Los agujeros están rellenos de un dieléctrico con permitividad ϵ_b . El sistema está parametrizado por el ancho W y alto H de los prismas y por el desplazamiento relativo ρ entre cada pareja. Notamos que este sistema no tiene simetría de reflexión a lo largo del plano, excepto para ciertos valores particulares del desplazamiento y que, aunque simétrica, su respuesta dieléctrica macroscópica es no hermitiana. Por lo tanto, los ejes principales en los que se diagonaliza el tensor dieléctrico de esta estructura son en general complejos, los modos propios respectivos corresponden a polarización elíptica y dependen en general de la frecuencia. Podemos aprovechar estas características del sistema para diseñar varios dispositivos ópticos. Como el cálculo reseñado en la sec. 2.6 es muy eficiente, podemos calcular espectros completos para cada una de las combinaciones de parámetros que surjan en una búsqueda automatizada del óptimo de cualquier propiedad deseada. [19] Por ejemplo, hemos hallado las combinaciones de parámetros que nos permiten obtener un máximo de dicroísmo circular, la diferencia entre la absortancia de la película cuando es iluminada con polarización circular derecha e izquierda, y situarlo en cualquier frecuencia deseada. [19] En el lado derecho de la fig. 16 mostramos fragmentos de los espectros de dicroísmo obtenidos, mostrando que podemos entonar su máximo a cualquier energía deseada en el espectro visible. Hacemos notar que el dicroísmo circular de los materiales naturales suele ser de apenas unas partes en mil, mientras

que aquí hemos encontrado señales de orden uno.

4.3. Magnetismo

La permitividad obtenida de acuerdo al procedimiento descrito en la sección 2.3 conduce a una permitividad que en general es *no-local*, es decir, $\epsilon(\mathbf{k}, \omega)$ depende explícitamente del vector de onda \mathbf{k} además de depender de la frecuencia ω . Una de las consecuencias esta no-localidad o *dispersión espacial*, es que la permitividad incluye información sobre la respuesta magnética del sistema. Para entender cómo un sistema no magnético adquiere propiedades magnéticas cuando se excita con un campo eléctrico cuya longitud de onda es del orden de las otras escalas de distancia del sistema, consideremos un cilindro metálico. Si iluminamos el cilindro con un campo cuya longitud de onda fuese el doble del diámetro del cilindro, entonces el campo eléctrico induciría corrientes en una dirección en la mitad del cilindro y en la dirección opuesta en la otra mitad, i.e., induciría una corriente que circularía alrededor del cilindro, generando un dipolo magnético de origen eléctrico.

Un procedimiento simple para extraer la respuesta magnética a partir de la dispersión espacial de la respuesta dieléctrica consiste en analizar la relación de dispersión de las ondas electromagnéticas en el medio no local,

$$k^2 = \epsilon_M(k, \omega) \frac{\omega^2}{c^2}, \quad (55)$$

donde por simplicidad ignoramos el carácter tensorial de ϵ y vectorial de k . Hacemos una expansión de Taylor

$$\epsilon_M(k, \omega) = \epsilon_M(0, \omega) + \frac{k^2}{2} \frac{\partial^2}{\partial k^2} \epsilon_M(0, \omega) + \dots, \quad (56)$$

respecto al número de onda para k 's pequeñas. El término lineal está ausente de esta expansión si el sistema es invariante frente a inversiones temporales. Sustitución en la ec. (55) conduce aproximadamente a

$$k^2 = \left(\epsilon_M(0, \omega) + \frac{k^2}{2} \frac{\partial^2}{\partial k^2} \epsilon_M(0, \omega) \right) \frac{\omega^2}{c^2}. \quad (57)$$

Despejando k^2 obtenemos

$$k^2 = \frac{\epsilon_M(0, \omega)}{1 - \frac{\omega^2}{2c^2} \frac{\partial^2}{\partial k^2} \epsilon_M(0, \omega)} \frac{\omega^2}{c^2}, \quad (58)$$

la cual podemos escribir como

$$k^2 = \epsilon_M(\omega) \mu_M(\omega) \frac{\omega^2}{c^2}, \quad (59)$$

donde definimos $\epsilon_M(\omega)$ como el límite local $\epsilon_M(k \rightarrow 0, \omega)$ de la permitividad no local, y donde identificamos la permeabilidad local

$$\mu_M(\omega) = \frac{1}{1 - \frac{\omega^2}{2c^2} \frac{\partial^2}{\partial k^2} \epsilon_M(0, \omega)}. \quad (60)$$

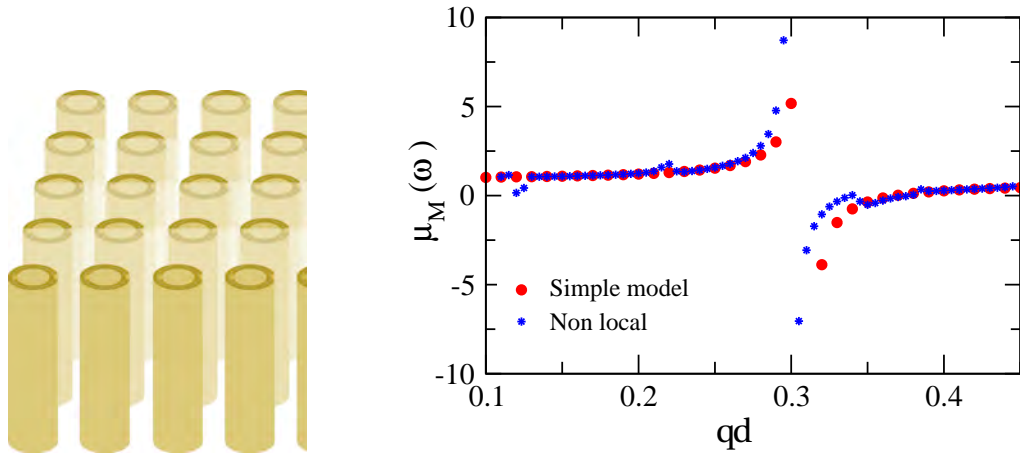


Figura 17: (Izquierda) Red cuadrada de parejas de cilindros metálicos cortados (split rings) anidados. Los cilindros están descritos por el modelo de Drude con frecuencia de plasma $\omega_p = 20c/d$, con d el parámetro de red. (Derecha) Permeabilidad magnética obtenida a partir de la dispersión espacial de la permitividad y a partir de un modelo simple para cilindros con paredes infinitesimales pero con la misma conductividad superficial. Los radios internos y externos de los anillos metálicos son $0.26d$, $0.32d$, $0.34d$, y $0.4d$ y están interrumpidos por brechas de tamaño $0.1d$. Se emplearon 350 pares de coeficientes de Haydock en una retícula de 401×401 píxeles.

En la fig. 17 mostramos un sistema formado por una red cuadrada, de anillos concéntricos truncados. Calculamos con *Photonic* la permeabilidad no local del sistema, y a partir de su dispersión espacial obtuvimos la permeabilidad magnética.[20] Del lado derecho mostramos el espectro de la permeabilidad $\mu_M(\omega)$ como función de la frecuencia normalizada $qd = \omega d/c$ para un sistema de anillos metálicos, descritos por una respuesta de Drude con frecuencia de plasma $\omega_p = 20c/d$, con d el parámetro de red. Los radios internos y externos de los anillos son $0.26d$, $0.32d$, $0.34d$, y $0.4d$ y están interrumpidos por brechas de tamaño $0.1d$. Los cálculos fueron realizados en una retícula de 401×401 píxeles y se emplearon 350 pares de coeficientes de Haydock. Como referencia, se muestran resultados de un cálculo simplificado para una red de anillos infinitamente delgados pero con una conductividad superficial que corresponde a los anillos sólidos. Hay un buen acuerdo entre ambos resultados. Se observa que alrededor de $qd = 0.3$ aparece una resonancia en la permeabilidad, arriba de la cual adquiere valores negativos, por lo cual este sistema podría emplearse para construir un metamaterial izquierdo.

4.4. Respuesta no lineal

Además de obtener la respuesta macroscópica, es posible obtener con una ligera extensión del formalismo presentado arriba el campo eléctrico microscópico en el seno de un metamaterial. El campo microscópico permite hacer cálculos de propiedades no lineales, tales y como la generación de segundo armónico proporcional al cuadrado del campo. En sistemas centrosimétricos, aquellos con simetría de inversión, no se pueden llevar a cabo procesos cuadráticos, en los que se absorben dos

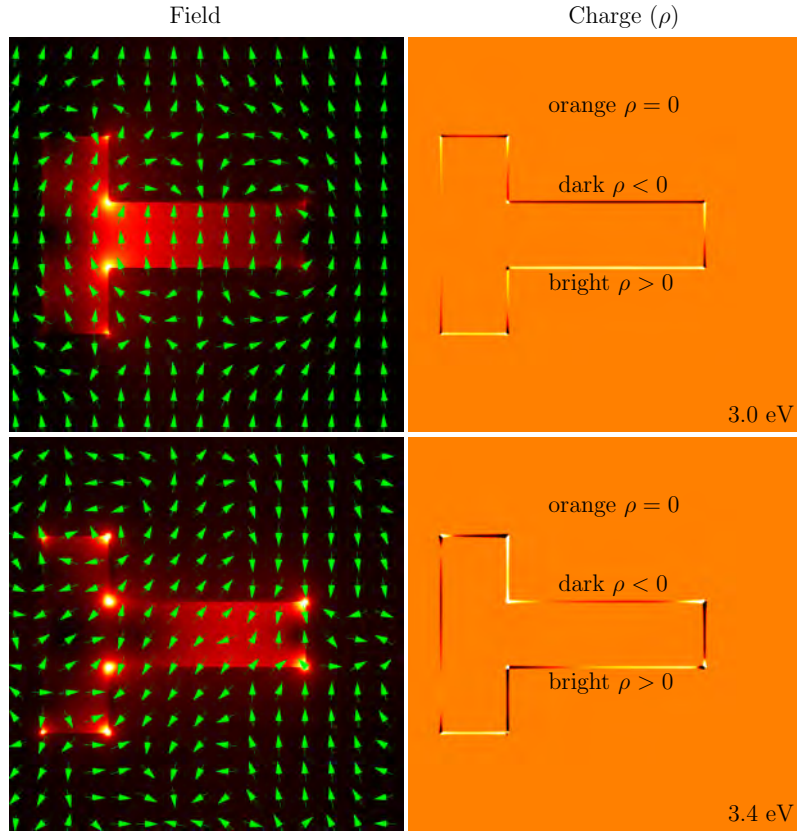


Figura 18: Magnitud y dirección del campo lineal microscópico (izquierda) y la densidad de carga inducida (derecha) en un metamaterial formado por una red de agujeros en forma de letra T en una matriz de Ag excitado por un campo polarizado en la dirección vertical para dos frecuencias resonantes. El carácter dipolar y cuadrupolar de las resonancias se advierte en los signos de la carga inducida y las direcciones del campo.

fotones y se emite un fotón con la suma de sus energías. En particular, no se pueden llevar a cabo procesos de generación de segundo armónico, en que dos fotones de frecuencia ω se combinen entre sí para dar lugar a un fotón de frecuencia 2ω . Sin embargo, en la vecindad de superficies estos procesos sí están permitidos, aunque para superficies centrosimétricas, las contribuciones de distintas partes opuestas de la superficie se cancelan mutuamente. Por esto es interesante calcular la respuesta no lineal de metamateriales formados por materiales centrosimétricos pero con geometrías no centrosimétricas. Como un ejemplo,[21] en la fig. 18 mostramos el campo lineal microscópico y la densidad de carga inducida en una red de agujeros no centrosimétricos con forma de letra T en el seno de una película de Ag. Debido a la no homogeneidad del campo lineal, en este sistema se induce una polarización cuadrática que oscila en el segundo armónico, como ilustra la fig. 19 para diversas direcciones de polarización del campo lineal macroscópico y diferentes frecuencias. Notamos que los patrones no lineales son simétricos, como el sistema, ante una reflexión $y \leftrightarrow -y$ cuando el campo \mathbf{E}_M apunta en las direcciones \hat{x} o \hat{y} , pero que esta simetría se pierde cuando el campo apunta en otras direcciones como la $\hat{x} + \hat{y}$.

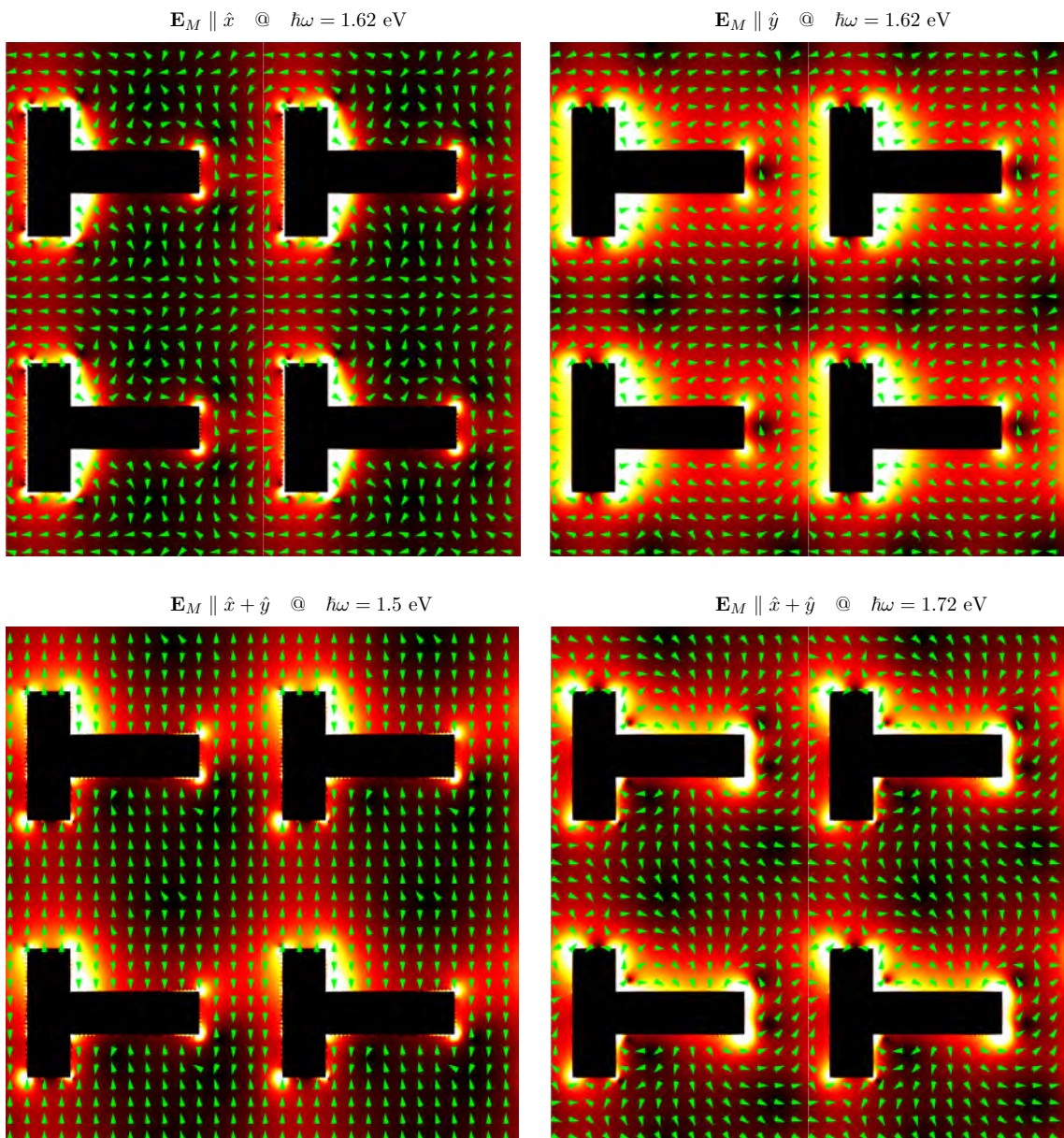


Figura 19: Magnitud y dirección de la polarización cuadrática no lineal inducida en el segundo armónico para el mismo sistema descrito en la fig. 18 para diversas frecuencias y diversas direcciones del campo eléctrico macroscópico lineal.

A partir de la polarización no lineal podemos calcular todas las componentes del tensor de susceptibilidad no lineal del metamaterial y podemos optimizarlo a través de los parámetros geométricos del sistema. En la fig. 20 mostramos las componentes $\chi_{ijk}(\omega, \omega; 2\omega)$ no nulas de la susceptibilidad cuadrática para la generación de segundo armónico en el sistema. La susceptibilidad típica de un material no centrosimétrico es del orden de $1/nea_B$, donde n es la densidad de número atómica, e la carga del electrón y a_B el radio de Bohr. Nuestros resultados muestran que un metamaterial hecho de componentes centrosimétricas pero con una geometría no centrosimétrica puede alcanzar en resonancia susceptibilidades del orden de $100/nea$, con a el parámetro de red. Por lo tanto, para materiales nanoestructurados, la respuesta no lineal de nuestros metamateriales puede ser competitiva con la de los materiales no lineales usuales.

5. Conclusiones

En este trabajo hemos presentado una introducción a los metamateriales y a algunas de sus múltiples propiedades, algunas exóticas, y aplicaciones. Luego desarrollamos una teoría basada en la identificación de operadores cuyo promedio tiene significado físico y a partir de los cuales podemos obtener las funciones respuesta macroscópicas del sistema y sus propiedades ópticas. También presentamos el método recursivo de Haydock, el cual aprovecha una analogía entre el cálculo de funciones respuesta macroscópicas y funciones de Green proyectadas para obtener algoritmos computacionales muy eficientes. Los diversos métodos desarrollados han sido implementados en paquetes computacionales modulares que han sido puestos en el dominio público. Como algunos ejemplos de su uso, presentamos el cálculo, diseño y optimización de propiedades ópticas lineales y no lineales como son la transmitancia extraordinaria, el dicroísmo lineal, el dicroísmo circular, las resonancias magnéticas y la generación de segundo armónico.

6. Agradecimientos

Este trabajo fue apoyado por DGAPA-UNAM mediante el proyecto IN111119. Parte del trabajo aquí reportado fue realizado en colaboración con varios colegas, incluyendo a Guillermo P. Ortiz, Bernardo S. Mendoza, José Samuel Pérez-Huerta, Lucila Juárez Reyes y Raksha Singla, así como sus grupos de trabajo y estudiantes.

Referencias

- [1] Virtual Institute for Artificial Electromagnetic Materials and Meta-materials. <https://www.metamorphose-vi.org/>
- [2] *An introduction to metamaterials and nanophotonics*, Constantin Simovski y Sergei Tretyakov, (Cambridge University Press, Cambridge, 2021), ISBN:9781108492645. <http://www.cambridge.org/9781108492645>

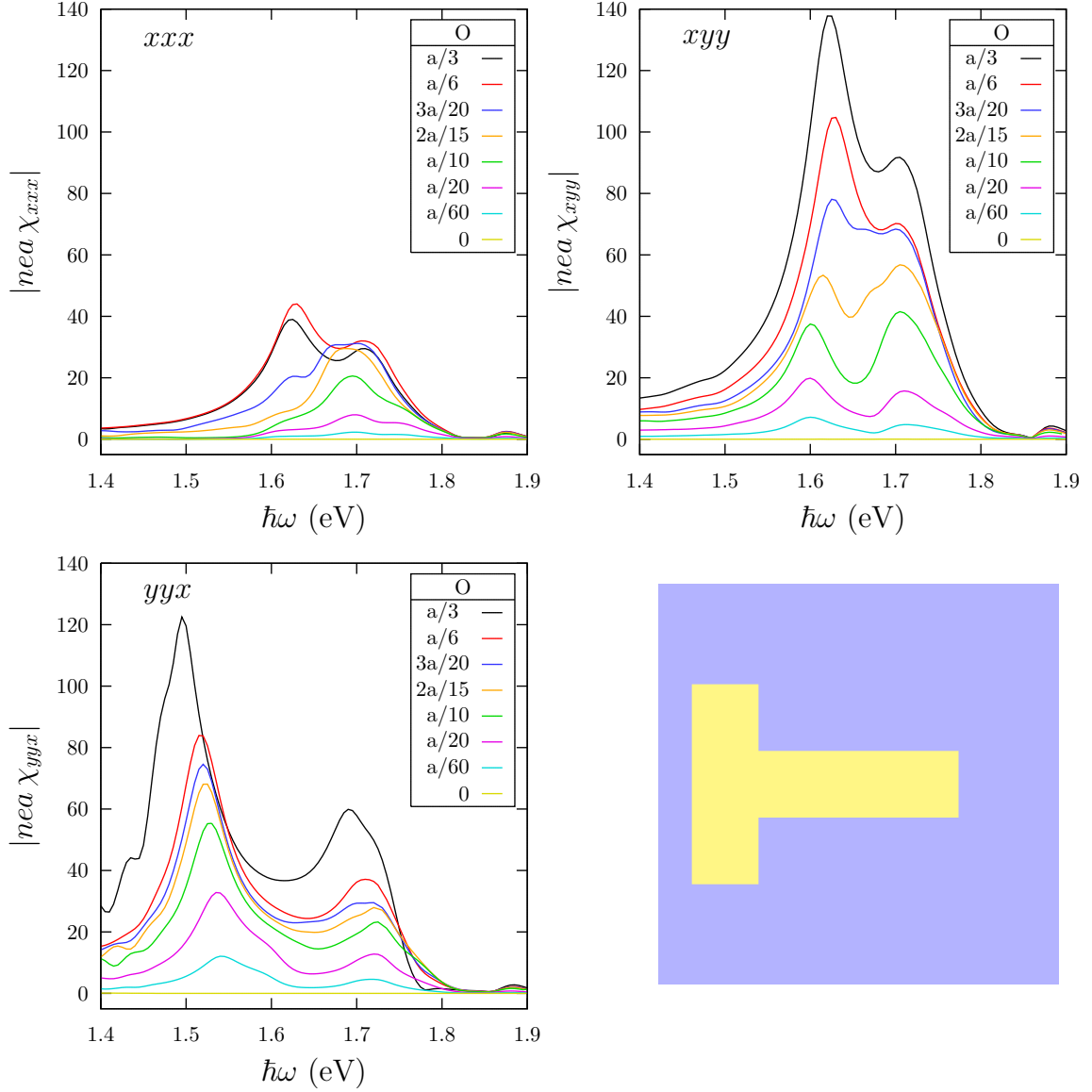


Figura 20: Distintas componentes de la susceptibilidad no lineal para generación de segundo armónico χ_{ijk} para el sistema descrito en la figura 18 como función de la energía de los fotones fundamentales. Distintas curvas corresponden a las contribuciones de zonas a distintas distancias de la superficie.

- [3] *Electromagnetic response of systems with spatial fluctuations. I. General formalism*, W. Luis Mochán y Rubén G. Barrera. Phys. Rev. B **32** 4984 (1985). doi: [10.1103/PhysRevB.32.4984](https://doi.org/10.1103/PhysRevB.32.4984).
- [4] *Plasmons*, W. Luis Mochán, en Reference Module in Materials Science and Materials Engineering, (Elsevier, Amsterdam, 2016) ed. por Saleem Hashmi. ISBN 9780128035818, doi:[10.1016/B978-0-12-803581-8.01192-9](https://doi.org/10.1016/B978-0-12-803581-8.01192-9).
- [5] *An investigation of the origin of the color of the Lycurgus cup by analytical transmission electron-microscopy*, Barber DJ y Freestone IC, Archaeometry **32** 33–45 (1990). doi:[10.1111/J.1475-4754.1990.TB01079.X](https://doi.org/10.1111/J.1475-4754.1990.TB01079.X).
- [6] *Theory of the optical properties of ionic crystal cubes*, Ronald Fuchs, Physical Review B **11** 1732, (1975). doi:[10.1103/PhysRevB.11.1732](https://doi.org/10.1103/PhysRevB.11.1732).
- [7] *Efficient homogenization procedure for the calculation of optical properties of 3D nanostructured composites*, W. Luis Mochán, Guillermo P. Ortiz y Bernardo S. Mendoza, Optics Express **18** 22119 (2010). doi:[10.1364/OE.18.022119](https://doi.org/10.1364/OE.18.022119).
- [8] *Macroscopic optical response and photonic bands*, José Samuel Pérez-Huerta, Guillermo P. Ortiz, Bernardo S. Mendoza y W. Luis Mochán, New Journal of Physics **15** 043037 (2013). doi:[10.1088/1367-2630/15/4/043037](https://doi.org/10.1088/1367-2630/15/4/043037).
- [9] *The Recursive Solution of the Schrödinger Equation*, R. Haydock, Solid State Physics **35** 215 (1980) ed. por H. Ehrenreich, F. Seitz y D. Turnbull.
- [10] W. Luis Mochán, Guillermo Ortiz, Bernardo S. Mendoza, José Samuel Pérez-Huerta, Lucila Juárez Reyes, Raksha Singla y Merlyn Jaqueline Juárez-Gutiérrez, *Photonic*, <https://metacpan.org/pod/Photonic>.
- [11] W. Luis Mochán, Guillermo Ortiz, Bernardo S. Mendoza, José Samuel Pérez-Huerta, Lucila Juárez Reyes, Raksha Singla y Merlyn Jaqueline Juárez-Gutiérrez, *Photonic*, <https://github.com/wlmb/Photonic>.
- [12] W. Luis Mochán, *Perl Benchmark* <https://wlmb.github.io/2021/10/03/pdl/>.
- [13] *Perl Data Language: Scientific Computing with Perl*. <http://pdl.perl.org/>.
- [14] *PDL* <https://metacpan.org/pod/PDL>.
- [15] *Perl Data Language (PDL)*. <https://github.com/PDLPorters/pdl>.
- [16] *Moose - A postmodern object system for Perl 5*. <https://metacpan.org/pod/Moose>.
- [17] [Ejemplo de uso de Photonic para el cálculo de las propiedades ópticas de una red de toroides.](#)

- [18] *Birefringent Nanostructured Composite Materials*, Bernardo S Mendoza y W. Luis Mochán, Physical Review B **85** 125418 (2012). [doi:10.1103/PhysRevB.85.125418](https://doi.org/10.1103/PhysRevB.85.125418).
- [19] *Tailored Optical Polarization in Nanostructured Metamaterials*, Bernardo S. Mendoza y W. Luis Mochán, Physical Review B **94** 195137 (2016). [doi:10.1103/PhysRevB.94.195137](https://doi.org/10.1103/PhysRevB.94.195137).
- [20] *Magnetic Response of Metamaterials*, Lucila Juárez-Reyes y W. Luis Mochán, Physica Status Solidi (b) **255** 1700495 (2018). [doi:10.1002/pssb.201700495](https://doi.org/10.1002/pssb.201700495).
- [21] *Second-Harmonic Generation in Nanostructured Metamaterials*, Ulises R. Meza, Bernardo S. Mendoza y W. Luis Mochán, Physical Review B **99** 125408 (2019). [doi:10.1103/PhysRevB.99.125408](https://doi.org/10.1103/PhysRevB.99.125408).

Notas de la XXVIII escuela de verano en Física
Semiconductores de baja dimensionalidad
Hugo Albero Lara García

Introducción

Un semiconductor es un sólido con una estructura cristalina periódica que, en el cero absoluto es un material eléctricamente aislante y un conductor a temperaturas altas que estén por debajo de su temperatura de fusión. En un semiconductor la función de onda de los electrones interactúa con el potencial periódico de la estructura cristalina, lo que resulta en una “propagación” de los niveles de energía en bandas, lo cual, también provoca la creación de regiones de energía prohibidas, en las cuales los electrones no pueden existir. Estas regiones se conocen como la banda prohibida o el bandgap del material. Las bandas de energía más importantes en una estructura cristalina son las bandas más cercanas al nivel de Fermi, recordemos que el nivel de Fermi es aquel nivel energético que tiene la probabilidad de $\frac{1}{2}$ de ser ocupado por un electrón. Estas bandas son conocidas como la banda de valencia (BV) y la banda de conducción (BC) y corresponde a las bandas energéticas más bajas de los estados electrónicos vacíos y a las bandas energéticas más altas de los estados electrónicos ocupados. La superposición de la banda de valencia y de conducción en los metales les confiere su característica propiedad de conductividad eléctrica, ya que los electrones pueden fluir fácilmente a estados en la banda de conducción. En cambio, para los aislantes, la banda de energía prohibida es suficientemente grande para evitar que los electrones “salten” de la banda de valencia a la banda de conducción y por lo tanto no fluyen a través del material.

En un semiconductor, esta brecha es relativamente pequeña y la conductividad va a depender de las posiciones de la BV y BC del material semiconductor específico. La corriente en un semiconductor puede fluir a medida que los electrones son excitados desde la banda de valencia a la banda de conducción ya sea por energía térmica o electromagnética. Cuando un electrón se excita desde la BV a la BC, se deja una vacante en la BV, conocido como hueco. Un hueco es una cuasi-partícula que tiene una masa efectiva determinada por la estructura de bandas del semiconductor y tiene una carga opuesta a la del electrón. Debido a sus cargas opuestas, estas partículas se pueden atraer entre sí creando otra cuasi-partícula electrón-hueco, a la cuál se le conoce como excitón.

Estructura de banda de los semiconductores

Debido a la estructura cristalina periódica y altamente simétrica de los semiconductores, se puede definir una celda unitaria que describe la red cristalina

completa repitiendo esta celda unitaria a través del semiconductor. Esto también permite definir una red recíproca que proporciona información de los ángulos relativos entre los planos y el tamaño de los planos en el cristal real del semiconductor. Cuando la celda unitaria se usa para definir el cristal, la zona de Brillouin se usa para definir la red recíproca. Por ejemplo, la primera zona de Brillouin para un cristal cúbico centrado en las caras se muestra en la figura 1. Los puntos resaltados en la red recíproca corresponden a los vectores de onda en el espacio real. En Γ (la zona central) el vector de onda y por lo tanto el momento son cero, y a medida que la distancia aumenta con respecto a la zona central, el momento aumenta.

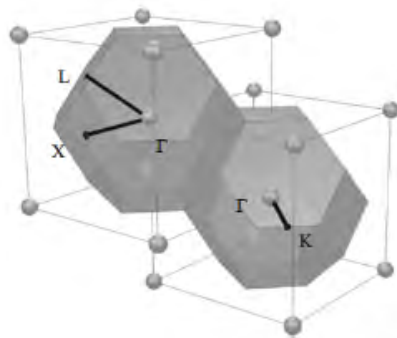


Figura 2. Primera zona de Brillouin para un cristal cúbico centrado en las caras.

Como la simetría del cristal cambia dependiendo del camino que se siga a través de la red, un electrón en movimiento tiene una energía que es dependiente de su dirección de movimiento. Los puntos de simetría resaltados en la figura 1 (L, K, X) actúan como una referencia conveniente cuando se discute la energía de un electrón como función del vector de onda o el momento. Graficar la energía de un electrón de esta manera se conoce como una relación de dispersión. La relación de dispersión para dos de los semiconductores más usados (Si y GaAs) se muestran en la figura 2. La forma compleja de estas relaciones de dispersión surge de la interacción de los portadores de carga con la red cristalina. Sin embargo, es claro donde se encuentra el máximo de la banda de valencia y el mínimo de la banda de conducción, la diferencia entre estos dos valores es el valor de la energía de la banda prohibida.

En este ejemplo, también se puede observar que el mínimo de la banda de conducción y el máximo de la banda de valencia no siempre están en el mismo vector de onda. En la relación de dispersión del GaAs, ambos puntos coinciden en el centro de la zona de Brillouin (Γ), lo que da como resultado un semiconductor con bandgap directo. Por lo contrario, en el caso del Si, las posiciones del mínimo de la BC y el máximo de la BV se encuentran en los puntos de simetría X y C, respectivamente. En este caso, el sistema tiene un bandgap indirecto.

Este atributo tiene implicaciones sobre las propiedades ópticas del material. Por ejemplo, si un fotón con energía mayor a la energía de la banda prohibida incide en el GaAs, un electrón podría excitarse de la banda de valencia a la banda de conducción, ya que el electrón no necesitaría ganar momento. En cambio, si se aplicará el mismo principio al Si, tuviese que haber un cambio en el momento para que el electrón se excitara a través de la banda prohibida. Es importante aclarar, que este no puede venir del fotón pues tiene un momento insignificante, el momento requerido se tiene que proporcionar a través de fonones (cuantos vibracionales de la red cristalina). Por lo tanto, para que ocurra una transición de un semiconductor con transición indirecta, un fotón o un fonón deben de ser absorbidos o emitidos. La absorción del Si se puede mejorar usando fotones de alta energía para excitar electrones a estados en la banda de conducción que no requieran el uso adicional de momento.

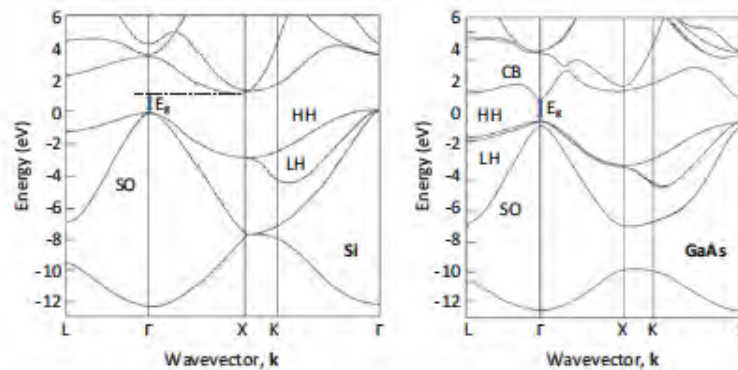


Figura 2. Estructura electrónica de bandas del Si y el GaAs.

La naturaleza directa o indirecta del semiconductor también tiene implicaciones en la emisión de luz. En semiconductores de brecha directa, un electrón que se ha excitado a la banda de conducción se puede recombinar radiativamente al caer de nuevo a la banda de valencia, esto provoca la emisión de un fotón para conservar la energía. Sin embargo, para los semiconductores de brecha indirecta, la recombinación sería un proceso de segundo orden, esta requiere la emisión de un fotón y la absorción o emisión de un fonón. Además, como el electrón permanece en su estado excitado durante un período mayor de tiempo, existe la posibilidad de recombinación no radiativa, a través de la ayuda de defectos, una recombinación Auger o una recombinación superficial. Por lo tanto, los semiconductores indirectos son malos emisores de luz y es la razón por la cual el silicio no se usa para aplicación relacionadas con esta propiedad.

Confinamiento cuántico y sus efectos

Las propiedades electrónicas y ópticas de un semiconductor se ven afectadas por el confinamiento cuántico. El confinamiento cuántico ocurre cuando el tamaño físico del semiconductor es comparable con la longitud de onda de De Broglie de los electrones y huecos. Si un electrón está en movimiento en un semiconductor con una temperatura T, la longitud de onda de De Broglie esta dada por:

$$\lambda = \frac{h}{\sqrt{m^*k_bT}}$$

Donde m^* es la masa efectiva del electrón o del hueco, h es la constante de Plank y k es la constante de Boltzman. Para un electrón en un cristal de GaAs (donde la masa efectiva es $m^*=0.067m_0$) a la temperatura ambiente (300 K), la longitud de onda de De Broglie es aproximadamente 42 nm. Esto significa que para nanoestructuras que están en el orden de nanómetros, el confinamiento cuántico gobernara las propiedades del material.

Por ejemplo, cuando una capa delgada de un material (decenas de nanómetros), es intercalado entre dos capas de otro material con un bandgap mayor, entonces se forma una capa que confina a los electrones y los huecos, esto se conoce como un pozo cuántico (figura 3). El confinamiento cuántico de los electrones y huecos conducen a estados de energía que ya no son continuos, sino discretos. La discretización de estos niveles de energía se puede predecir mediante el modelo de "partícula en una caja". En su forma más sencilla, la ecuación de Schrödinger independiente del tiempo es:

$$-\frac{\hbar^2}{2m^*} \frac{d^2\psi_n(z)}{dz^2} + V(z)\psi_n(z) = E_n\psi_n(z)$$

Donde $V(z)$ es el potencial visto por el electrón o el hueco a lo largo de z , m^* es la masa efectiva, ψ_n y E_n son la función y la energía propias para cada valor de n .

Si consideramos una barrera infinita ($V(z)=0$ en el pozo y $V(z)=\infty$ fuera del pozo), donde la función de onda no puede existir fuera del pozo, la solución se vuelve particularmente simple:

$$E_n = \frac{\hbar^2}{2m^*} \left(\frac{n\pi}{L}\right)^2 \quad \psi_n = A \sin\left(\frac{n\pi z}{L}\right)$$

Donde L es el ancho del pozo cuántico. Es importante notar que los niveles de energía están relacionados con el ancho del pozo de potencial, por lo cual el bandgap de un semiconductor, así como muchas de sus propiedades optoelectrónicas se pueden ajustar cambiando deliberadamente el ancho del pozo de potencial, lo que lleva a una metodología relativamente simple para diseñar un

semiconductor con propiedades ideales, sintonizándolo por ejemplo para emitir en el infrarrojo.

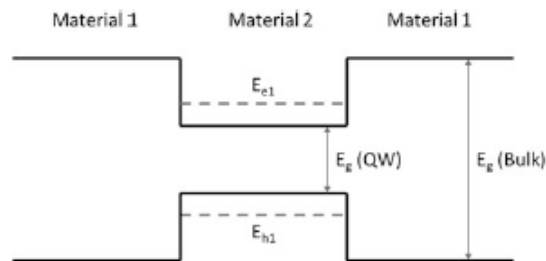


Figura 3. Materiales con diferente valor de bandgap que causan un confinamiento cuántico formando un pozo cuántico.

Los primeros tres valores de energía que surgen de estas ecuaciones se muestran en la Figura 4. Se puede ver que el primer nivel de energía confinado se encuentra por encima del borde de la banda. Esta es una observación interesante y se conoce como la energía del punto cero o la energía del confinamiento. Surge como resultado del principio de incertidumbre de Heisenberg. A medida que la partícula se ve limitada a una región finita del espacio, la variación en su posición alcanza un límite superior. Por lo tanto, el momento de las partículas no puede ser cero y debe poseer alguna energía finita para asegurar que no viole el principio de incertidumbre. Esta energía varía con el ancho del pozo, ya que a un mayor confinamiento hay una menor variabilidad en el espacio, por lo cual se observa una energía mayor del punto cero.

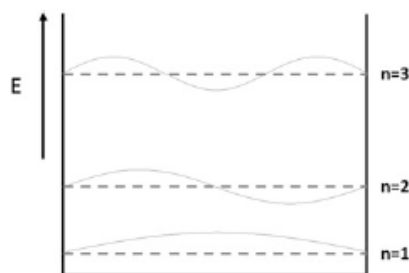


Figura 4. Funciones propias de los primeros tres niveles de energía asumiendo una partícula en una caja.

En un pozo cuántico, los electrones y los huecos sólo están confinados en una dimensión, en las dos dimensiones restantes existe un continuo de estados electrónicos. Se puede confinar a los electrones y huecos en 2 dimensiones y 3 dimensiones. Las nanoestructuras que confinan a los electrones en dos dimensiones se les conoce como 1D y pueden ser nanotubos, nanoalambres,

nanobarras, entre otras. Mientras que las nanoestructuras que los confinan en tres dimensiones a los electrones se conocen como puntos cuánticos (0D).

La densidad de estados para los materiales 3D, 2D, 1D y 0D se ilustran en la figura 5. Las funciones y los valores propios son diferentes en cada sistema, sin embargo, los siguientes efectos son generales:

- Con un confinamiento dimensional más alto, se observa un mayor desplazamiento al azul en el borde de la banda del semiconductor, esto se debe al principio de incertidumbre de Heisenberg, mencionado anteriormente.
- Como los electrones y los huecos están confinados en una región más pequeña del espacio, la superposición de la función de onda aumenta, por lo cual la probabilidad de la recombinación radiativa aumenta.
- La densidad de estados muestra una mayor discretización ya que las partículas tienen menor libertad de movimiento, hasta que finalmente en los puntos cuánticos se observan niveles de energía completamente discretos, estos niveles de energía son similares a los átomos, por esta razón los puntos cuánticos a veces se les denomina átomos artificiales.

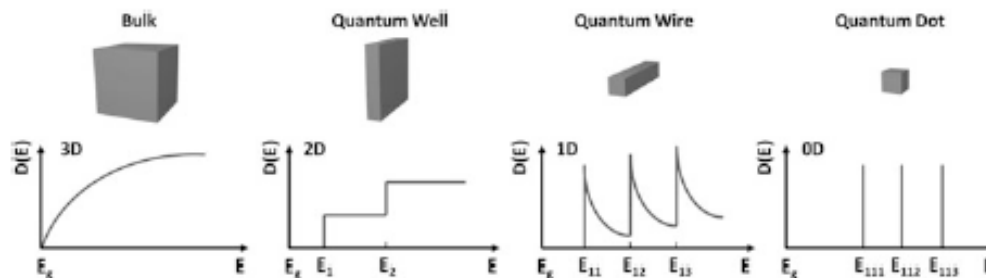


Figura 5. Densidad de estados de semiconductores variando las dimensiones confinadas.

Bibliografía

- M. Cardona, F.H. Pollak, Energy-band structure of germanium and silicon: the k_p method. *Phys. Rev.* 142, 530 (1966)
- H. Ehrenreich, Band structure and electron transport of GaAs. *Phys. Rev.* 120, 1951 (1960)
- J.F. Prins, Fermi-Dirac statistics and the nature of the compensating donors in boron-doped diamond layers. *Phys. Rev. B.* 39, 3764 (1989)
- G.H. Wannier, On the energy band structure of insulators. *Phys. Rev.* 76, 438 (1949)
- C.S. Hung, V.A. Johnson, Resistivity of semiconductors containing both acceptors and donors. *Phys. Rev.* 79, 535 (1950)
- L.P. Bouckaert et al., Theory of Brillouin zones and symmetry properties of wave functions in crystals. *Phys. Rev.* 50, 58 (1936)
- P.Y. Yu, M. Cardona, *Fundamentals of Semiconductors* (Springer, Berlin, 2010)
- G.H. Parker, C.A. Mead, Energy-momentum relationship of InAs. *Phys. Rev. Lett.* 21, 605 (1968)

9. W.G. Spitzer, H.Y. Fan, Determination of optical constants and carrier effective mass of semiconductors. *Phys. Rev.* 106, 882 (1957)
10. R.N. Dexter, B. Lax, Effective masses of holes in silicon. *Phys. Rev.* 96, 223 (1954)
11. P.T. Landsberg, M.J. Adams, Radiative and Auger processes in semiconductors. *J. Lumin.* 7, 3 (1973)
12. D.J. Fitzgerald, A.S. Grove, Surface recombination in semiconductors, in *Electron Devices Meeting* (1967)
13. R. Hill, Energy-gap variations in semiconductor alloys. *J. Phys. C Solid State Phys.* 7, 521 (1974)
14. C.B. Roxlo et al., Evidence for lattice-mismatch—induced defects in amorphous semiconductor heterostructures. *Phys. Rev. Lett.* 52, 1994 (1984)
15. M. Fox, *Optical Properties of Solids* (Oxford University Press, 2010)
16. J. Roberts, *Using Imperfect Semiconductor Systems for Unique Identification*, Springer Theses.

Simulaciones Cuánticas en Sistemas Electrónicos Sintéticos

Roberto de J. León-Montiel (roberto.leon@nucleares.unam.mx)
Instituto de Ciencias Nucleares, ICN-UNAM

El objetivo de estas notas, preparadas en el marco de la XXVIII Escuela de Verano en Física de la UNAM, es proporcionar una descripción detallada de cómo construir una analogía, teórica-experimental, entre el transporte cuántico de energía (en la base de una sola excitación) y la dinámica de circuitos electrónicos acoplados.

1. Modelo cuántico: N qubits acoplados

Para establecer las condiciones en las que podemos construir un simulador clásico de procesos de transporte cuántico, comenzamos describiendo la dinámica de una sola excitación en un sistema cuántico compuesto por N sistemas de dos niveles (o qubits) acoplados. El Hamiltoniano para este tipo de sistemas está dado por

$$\hat{H} = \sum_{n=1}^N \varepsilon_n |n\rangle \langle n| + \sum_{n \neq m}^N J_{nm} |n\rangle \langle m|, \quad (1)$$

donde $|n\rangle$ nos indica que la excitación se encuentra en el n -ésimo sitio. Las energías de excitación para cada sitio y los acoplamientos entre ellos están dados por los coeficientes ε_n y J_{nm} , respectivamente.

Para explorar la dinámica de la energía en la red de sitios acoplados, es conveniente expresar la función de onda del sistema en la base de sitios, i.e., $|\psi(t)\rangle = \sum_n c_n(t) |n\rangle$. Sustituyendo esta función de onda en la ecuación de Schrodinger, $i\partial_t |\psi(t)\rangle = \hat{H} |\psi(t)\rangle$, obtenemos

$$i \frac{dc_n}{dt} = \varepsilon_n c_n + \sum_{n \neq m}^N J_{nm} c_m. \quad (2)$$

Note que este resultado muestra el conjunto de ecuaciones diferenciales acopladas que describen la dinámica de las amplitudes de probabilidad de ocupación $c_n(t)$ para cada sitio. Es fácil demostrar que, en el límite de acoplamiento débil ($J_{nm} \ll \varepsilon_n$), la derivada temporal de la ecuación (2) toma la forma [1, 2, 3, 4]

$$\frac{d^2 c_n}{dt^2} = -\varepsilon_n^2 c_n - \varepsilon_n \sum_{n \neq m}^N 2J_{nm} c_m. \quad (3)$$

Como se mostrará a continuación, la importancia de este resultado radica en el hecho de que nos permitirá establecer una conexión directa entre las amplitudes de probabilidad cuánticas c_n , y las señales de voltaje presentes en una red de osciladores electrónicos.

2. Modelo clásico: Circuitos electrónicos acoplados

Para establecer un modelo electrónico equivalente a la ecuación (3), consideremos un arreglo de N osciladores RLC (donde R corresponde a resistencia, L a inductancia y C a capacitancia) acoplados inductivamente. La Figura 1 muestra una representación esquemática del arreglo de osciladores electrónicos.

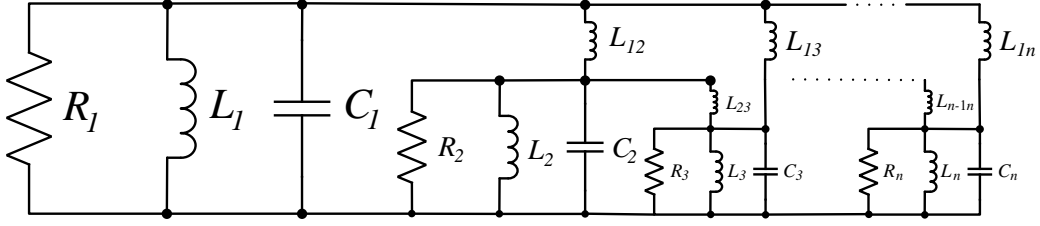


Figura 1: Representación esquemática de N osciladores RLC acoplados inductivamente.

Podemos entonces hacer uso de las leyes de Kirchhoff para describir la dinámica de los voltajes $V_n(t)$ en los capacitores C_n , las corrientes $I_n(t)$ en los inductores de cada oscilador L_n , y las corrientes $I_{nm}(t)$ en los inductores de acoplamiento L_{nm} , a través del siguiente conjunto de ecuaciones

$$\frac{dV_n}{dt} = \frac{1}{C_n} \left[-\frac{V_n}{R_n} - I_n - \sum_{j=n+1}^N I_{nj} + \sum_{j=1}^{j<n} I_{jn} \right], \quad (4)$$

$$\frac{dI_n}{dt} = \frac{1}{L_n} V_n, \quad (5)$$

$$\frac{dI_{nm}}{dt} = \frac{1}{L_{nm}} (V_n - V_m). \quad (6)$$

Note que sustituyendo las ecuaciones (5) y (6) en la derivada respecto al tiempo de la ecuación (4), podemos construir una ecuación de segundas derivadas respecto al tiempo dada por

$$\frac{d^2V_n}{dt^2} = \frac{1}{C_n} \left[-\frac{1}{R_n} \frac{dV_n}{dt} - \frac{V_n}{L_n} - \sum_{j=n+1}^N \frac{V_n - V_j}{L_{nj}} + \sum_{j=1}^{j<n} \frac{V_j - V_n}{L_{jn}} \right]. \quad (7)$$

Este resultado es una de las principales contribuciones de estas notas, pues nos permite establecer relaciones entre los parámetros del circuito electrónico y las energías de sitio y acoplamientos del sistema cuántico. A continuación se presenta a detalle la derivación del mapeo entre los parámetros cuánticos y los de circuitos electrónicos.

2.1. Sistema de 2 sitios

Para construir el mapeo entre el sistema de qubits acoplados y circuitos electrónicos, consideremos el caso más simple de dos sitios acoplados. Por simplicidad, asumimos inicialmente que no hay pérdidas de energía en el circuito, es decir, tomamos el límite donde $R_n \rightarrow \infty$. En esta situación, las ecuaciones para dos circuitos acoplados toman la forma

$$\frac{d^2V_1}{dt^2} = -\left(\frac{1}{L_1} + \frac{1}{L_{12}}\right) \frac{1}{C_1} V_1 + \frac{1}{L_{12}C_1} V_2, \quad (8)$$

$$\frac{d^2V_2}{dt^2} = -\left(\frac{1}{L_2} + \frac{1}{L_{12}}\right) \frac{1}{C_2} V_2 + \frac{1}{L_{12}C_2} V_1. \quad (9)$$

$$(10)$$

De igual forma, tomando $n = 2$ en la ecuación (3), encontramos

$$\frac{d^2c_1}{dt^2} = -\varepsilon_1^2 c_1 - \varepsilon_1 2J_{12} c_2, \quad (11)$$

$$\frac{d^2c_2}{dt^2} = -\varepsilon_2^2 c_2 - \varepsilon_2 2J_{21} c_1. \quad (12)$$

Comparando ambos conjuntos de ecuaciones, podemos establecer las siguientes relaciones

$$\varepsilon_1 = \sqrt{\left(\frac{1}{L_1} + \frac{1}{L_{12}}\right) \frac{1}{C_1}} \quad ; \quad \varepsilon_2 = \sqrt{\left(\frac{1}{L_2} + \frac{1}{L_{12}}\right) \frac{1}{C_2}}, \quad (13)$$

$$J_{12} = -\frac{1}{2\varepsilon_1 L_{12} C_1} \quad ; \quad J_{21} = -\frac{1}{2\varepsilon_2 L_{12} C_2}. \quad (14)$$

Para construir una generalización de este resultado, a continuación consideramos el caso con un oscilador adicional, es decir, una red de tres osciladores o trímico.

2.2. Sistema de 3 sitios

De forma totalmente análoga al caso anterior, podemos escribir las ecuaciones para el circuito de osciladores acoplados como

$$\frac{d^2 V_1}{dt^2} = -\left(\frac{1}{L_1} + \frac{1}{L_{12}} + \frac{1}{L_{13}}\right) \frac{1}{C_1} V_1 + \frac{1}{L_{12} C_1} V_2 + \frac{1}{L_{13} C_1} V_3, \quad (15)$$

$$\frac{d^2 V_2}{dt^2} = -\left(\frac{1}{L_2} + \frac{1}{L_{12}} + \frac{1}{L_{23}}\right) \frac{1}{C_2} V_2 + \frac{1}{L_{12} C_2} V_1 + \frac{1}{L_{23} C_2} V_3, \quad (16)$$

$$\frac{d^2 V_3}{dt^2} = -\left(\frac{1}{L_3} + \frac{1}{L_{13}} + \frac{1}{L_{23}}\right) \frac{1}{C_3} V_3 + \frac{1}{L_{13} C_3} V_1 + \frac{1}{L_{23} C_3} V_2. \quad (17)$$

Mientras que las ecuaciones del sistema cuántico toman la forma

$$\frac{d^2 c_1}{dt^2} = -\varepsilon_1^2 c_1 - \varepsilon_1 2J_{12} c_2 - \varepsilon_1 2J_{13} c_3, \quad (18)$$

$$\frac{d^2 c_2}{dt^2} = -\varepsilon_2^2 c_2 - \varepsilon_2 2J_{21} c_1 - \varepsilon_2 2J_{23} c_3, \quad (19)$$

$$\frac{d^2 c_3}{dt^2} = -\varepsilon_3^2 c_3 - \varepsilon_3 2J_{31} c_1 - \varepsilon_3 2J_{32} c_2. \quad (20)$$

Comparando ambos conjuntos de ecuaciones, encontramos las relaciones

$$\varepsilon_1 = \sqrt{\left(\frac{1}{L_1} + \frac{1}{L_{12}} + \frac{1}{L_{13}}\right) \frac{1}{C_1}},$$

$$\varepsilon_2 = \sqrt{\left(\frac{1}{L_2} + \frac{1}{L_{12}} + \frac{1}{L_{23}}\right) \frac{1}{C_2}}, \quad (21)$$

$$\varepsilon_3 = \sqrt{\left(\frac{1}{L_3} + \frac{1}{L_{13}} + \frac{1}{L_{23}}\right) \frac{1}{C_3}},$$

$$J_{12} = -\frac{1}{2\varepsilon_1 L_{12} C_1} \quad ; \quad J_{21} = -\frac{1}{2\varepsilon_2 L_{12} C_2}, \quad (22)$$

$$J_{13} = -\frac{1}{2\varepsilon_1 L_{13} C_1} \quad ; \quad J_{31} = -\frac{1}{2\varepsilon_3 L_{13} C_3}, \quad (23)$$

$$J_{23} = -\frac{1}{2\varepsilon_2 L_{23} C_2} \quad ; \quad J_{32} = -\frac{1}{2\varepsilon_3 L_{23} C_3}. \quad (24)$$

Usando los casos anteriores, podemos entonces deducir que el mapeo entre un sistema de N osciladores electrónicos acoplados y N qubits interactuantes está dado por los siguientes conjuntos de ecuaciones

$$\varepsilon_n = \sqrt{\left(\frac{1}{L_n} + \sum_{m \neq n}^N \frac{1}{L_{nm}}\right) \frac{1}{C_n}}, \quad (25)$$

$$J_{nm} = -\frac{1}{2\varepsilon_n L_{nm} C_n}. \quad (26)$$

Las ecuaciones (25) y (26) constituyen el principal resultado de estas notas, pues nos permiten identificar los parámetros electrónicos necesarios, para simular el sistema cuántico descrito por la ecuación (2). Con este mapeo, podemos implementar diferentes protocolos cuánticos en nuestra plataforma electrónica. En el seminario discutiremos algunos ejemplos de sistemas cuánticos que hemos implementado en el laboratorio de micro y nanofotónica del ICN-UNAM, y que han sido reportados recientemente en la Ref. [5].

Antes de concluir estas notas, es importante mencionar que en las realizaciones experimentales, la presencia de pérdidas de energía (debido a resistencias parásitas en los circuitos) no pueden ser despreciadas, por lo que es necesario introducir estos en efectos en el modelo cuántico. En la siguiente sección se discute cómo podemos realizar esto.

3. Análisis de pérdidas de energía

En esta sección presentamos una descripción detallada de cómo las pérdidas de energía, presentes en nuestra plataforma electrónica de simulaciones cuánticas, pueden ser incluidas en el modelo cuántico de transporte. Consideremos la energía contenida en el circuito electrónico, la cual está dada por

$$Q_{cl}(t) = \frac{1}{2} \sum_{m \neq n} C_n V_n^2 + L_n I_n^2 + L_{nm} I_{nm}^2. \quad (27)$$

Usando esta expresión, junto con la ecuaciones (4), (5) y (6), podemos encontrar que en presencia de pérdidas (o resistencia), la energía total del sistema decae exponencialmente [3]. Es interesante notar que en el modelo cuántico, una cantidad que sigue el mismo comportamiento en la presencia de pérdidas, es la traza de la matriz de densidad que describe al sistema cuántico. Podemos entonces definir una medida de “energía” del sistema cuántico, escribiendo

$$Q_q = \sum_n |c_n|^2. \quad (28)$$

Es bien conocido que para un sistema cuántico cerrado, la traza de la matriz de densidad se conserva. Sin embargo, en presencia de fenómenos disipativos, la traza decae exponencialmente [6]. En vista de esto, la forma más simple de introducir un proceso disipativo en un sistema cuántico es incluir un término no-Hermitiano en el Hamiltoniano del sistema cerrado [ecuación (1)]. Este término tiene la forma [7]

$$H_{\text{loss}} = -\frac{i}{2} \sum_n \Gamma_n |n\rangle \langle n|, \quad (29)$$

con Γ_n describiendo la razón de energía disipada al ambiente que rodea al sistema cuántico.

La Figura 2 muestra la evolución temporal de la energía del circuito eléctrico (en presencia de resistencia) y la traza de la matriz de densidad del sistema cuántico abierto. Note que ambas curvas siguen el mismo comportamiento exponencial (descrito por el ajuste teórico dado por la línea verde punteada). Esto nos permite establecer la relación

$$\Gamma_n = \frac{1}{RC}, \quad (30)$$

donde $R = R_n$ y $C = C_n$ representan las resistencias y las capacitancias en cada uno de los osciladores que constituyen la red electrónica. Este importante resultado es lo que nos permite incluir los efectos de las resistencias electrónicas en el modelo de transporte cuántico de energía.

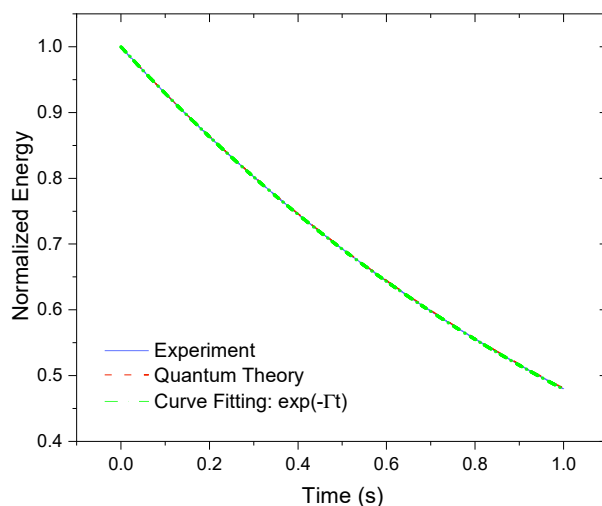


Figura 2: Evolución temporal de la energía normalizada del circuito electrónico (línea sólida) y del modelo cuántico abierto (línea punteada). La línea punto-rama muestra un ajuste exponencial de la curva. Para más detalles, ver la Ref. [5], que es la fuente de donde fue tomada esta gráfica.

Referencias

- [1] J. S. Briggs and A. Eisfeld, *Phys. Rev. E* **83**, 051911 (2011).
- [2] R. de J. León-Montiel and J. P. Torres, *Phys. Rev. Lett.* **110**, 218101 (2013).
- [3] R. de J. León-Montiel, M. A. Quiroz-Juárez, R. Quintero-Torres, J. L. Domínguez-Juárez, H. M. Moya-Cessa, J. P. Torres, and J. L. Aragón, *Sci. Rep.* **5**, 17339 (2015).
- [4] R. de J. León-Montiel, M. A. Quiroz-Juárez, J. L. Domínguez-Juárez, R. Quintero-Torres, J. L. Aragón, A. K. Harter, and Y. N. Joglekar, *Commun. Phys.* **1**, 88 (2018).
- [5] M. A. Quiroz-Juárez, C. You, J. Carrillo-Martínez, D. Montiel-Álvarez, J. L. Aragón, O. S. Magaña-Loaiza, and R. de J. León-Montiel, *Phys. Rev. Research* **3**, 013010 (2021).
- [6] H.-P. Breuer and F. Petruccione, *The theory of open quantum systems* (Oxford University Press, Oxford, 2002).
- [7] P. Rebentrost, M. Mohseni, I. Kassal, S. Lloyd, and A. Aspuru-Guzik, *New J. Phys.* **11**, 033003 (2009).

FUNDAMENTOS DE ELECTROMAGNETISMO Y DE MECANICA CUANTICA (I-III)

Eugenio Ley Koo, Instituto de Fisica UNAM

I. INTRODUCCION

Esta contribucion para CIENCIAS esta basada en el Ciclo Semestral de doce Conferencias con el mismo titulo, impartido en la Facultad de Ciencias en Febrero-Marzo y Agosto-Septiembre de 2020. Una contribucion del autor en la misma revista "La Hipotesis Cuantica de la Luz" (2005) se concentro en explicar el desarrollo de la mecanica cuantica y sus aplicaciones a lo largo de su primer siglo. En la presente contribucion se subraya la importancia de reconocer, revisar y conectar los fundamentos comunes del electromagnetismo y de la mecanica cuantica. Efectivamente, las doce conferencias se dividen en tres tercios con elementos conceptuales comunes cada uno, y que aclaran dudas comunes en una u otra disciplina cuando se plantean y analizan conjuntamente. Especificamente, en el primer tercio se consideran sucesivamente 1) El Principio de Superposicion, 2) El Problema de Eigenvalores, 3) El Significado Clasico y Cuantico de los Potenciales Electromagneticos y 4) Soluciones Lineales de las Ecuaciones Dinamicas y Propiedades Dinamicas Cuadraticas en esas soluciones. El tercio intermedio esta enfocado en la radiacion electromagnetica y sus interacciones con la materia: 5) Simetrias de las ecuaciones de Maxwell, 6) Nuevas Formas de Luz, 7) Nuevas formas de materia y de procesos, 8) Del Teorema de Helmholtz al Formalismo de Debye. El tercio final esta dedicado a las interpretaciones de la mecanica cuantica en base a experimentos pensados versus realizados: 9) Einstein, Bohr, Schroedinger 1935, 10) Fotones enredados en Aspect 1982, telefonia suiza 1995, y satelite y tres observatorios 2018, 11) Gatitos de Schroedinger, Saltos cuanticos en un solo ion y en un solo foton, Wineland y Haroche 2012, 12) Coherencia y Decoherencia, Zurek 1991.

Consecuentemente, la presentacion de esta contribucion se realiza en cinco secciones adicionales en que se analizan sucesivamente: Seccion II, Ecuaciones de Maxwell y de Schroedinger, Radiacion Electromagnetica y Materia, y sus Interacciones, las conexiones generales entre los tres tercios del ciclo identificando los problemas especificos de cada uno, sus metodologias comunes y extensiones, y con ilustraciones de nuevas soluciones; Seccion III. Fundamentos Comunes de Electromagnetismo y Mecanica Cuantica, detalles ilustrativos de principios, metodologias, conceptos y estructuras comunes en ambas areas; Seccion IV. Nuevas formas de radiacion electromagnetica, de materia y de sus interacciones; Seccion V. Interpretaciones de la Mecanica Cuantica y experimentos que las han puesto a prueba; y VI. Discusion.

II Ecuaciones de Maxwell y de Schroedinger, Radiacion Electromagnetica y Materia, y sus Interacciones.

El lector puede apreciar las diferencias entre los titulos de las conferencias en el ciclo semestral y en el primer parrafo de la Introduccion, debido a que en esta version escrita el objetivo es destacar lo comun y lo diferente en los fenomenos electromagneticos y en los fenomenos cuanticos. Tambien se deben senalar los titulos representativos de cada tercio del Ciclo, y de su conjunto en esta seccion. Aqui el objetivo es dar una descripcion breve de cada tercio, y subrayar las conexiones sucesivas entre ellos tomando en cuenta que el primer tercio contiene el formalismo matematico de cada area; el segundo tercio discute la fisica de nuevas formas de luz o de radiacion electromagnetica, de materia y de sus interacciones, como soluciones de las ecuaciones del primer tercio; y el tercer tercio se enfoca en experimentos con formas elementales de materia y de luz para poner a prueba las interpretaciones de la mecanica cuantica, comparando las predicciones de la mecanica cuantica con las de la teoria de variables escondidas de Bell, y afirmaciones de Einstein y Schroedinger.

En electromagnetismo y en mecánica cuántica se reconocen la validez de principios de superposición; de el problema de eigenvalores que tiene asociados efectos de cuantización; el uso de los potenciales electromagnéticos, escalar y vectorial, con significados diferentes en cada área; y el uso de campos de fuerza y de funciones de onda, respectivamente, y de formas cuadráticas de los mismos asociadas a las propiedades dinámicas de los sistemas respectivos. En estos temas se formulan las siguientes preguntas sucesivas: Existe una conexión entre los principios de superposición de electromagnetismo y de mecánica cuántica? Schrodinger tituló sus trabajos Cuantización como Problema de Eigenvalores, ¿Cuál es la cuantización en electromagnetismo? Aharonov y Bohm plantearon la pregunta sobre el significado cuántico de los potenciales electromagnéticos; el Autor agrega la pregunta sobre el significado clásico, y subraya la conexión entre ellos. Dyson argumentó que el nivel lineal en el uso de campos y funciones de onda no es observable, y el nivel cuadrático sí lo es; el Autor señala que el carácter ondulatorio de la luz y de la materia se observa directamente en fenómenos de interferencia y difracción en el primer nivel. Las respuestas a estas preguntas se encuentran al comparar las ecuaciones de Maxwell y de Schrodinger, y sus soluciones para la radiación electromagnética y la materia.

Las leyes de la física tienen un sustento observacional y experimental, y también formulaciones matemáticas. Las últimas permiten el análisis de sistemas físicos específicos y describir sus comportamientos en situaciones conocidas o nuevas: para las primeras se pueden formular explicaciones de los comportamientos experimentalmente observados; para las segundas se pueden formular predicciones que conducen a planear y realizar nuevos experimentos para ponerlas a prueba. El lector puede ver ejemplos de este tipo en el ciclo de conferencias Historias Clásicas y Modernas de Física y Matemáticas del Autor en TV Ciencias, y en el artículo antes mencionado de Ciencias en que se ilustra cómo los intentos de entender los espectros característicos de las radiaciones emitidas o absorbidas por los elementos químicos condujeron a la formulación de la mecánica cuántica y su desarrollo.

El segundo tercio de las conferencias ilustra ejemplos más recientes de las Historias anteriores. Específicamente se identifican las simetrías de las ecuaciones de Maxwell y sus consecuencias; se explica la conexión entre los procesos de emisión espontánea e inducida y de absorción inducida de la luz propuestos por Einstein en 1916 y la invención del maser en 1955 y el láser en 1959, así como la conexión entre la solución exacta de las ecuaciones de Maxwell en coordenadas cilíndricas circulares, elípticas y parabólicas y la implementación experimental de los haces adifraccionales respectivos en Rochester 1987, INAOE 2000, e Instituto Tecnológico de Monterrey 2003; la interacción entre la luz y la materia ocurre a través de los procesos de polarización eléctrica y de magnetización según los libros, pero recientemente se han reportado observaciones en metamateriales y nanomateriales de un tercer proceso, llamado toroidización, anticipado desde 1957 por Zeldovich en su hipótesis del anapolo magnético; del teorema de Helmholtz al formalismo de Debye permite distinguir la diferencia entre los tres procesos anteriores en base de las direccionalidades de corrientes: longitudinal y transversal en el teorema, y la distinción de dos transversalidades en el formalismo, toroidal con líneas de corriente en círculos paralelos y poloidal con líneas de corriente cerradas en planos meridianos, y en ambos casos el campo de intensidad eléctrica comparte la direccionalidad de las corrientes y el campo de inducción magnética es transversal a las corrientes por la ley de Ampere.

El tercio final de las conferencias ilustra ejemplos de Historias aún más recientes con la condición adicional de que se consideran las interacciones en el nivel elemental de la luz y de la materia. Las interpretaciones de la mecánica de Einstein, Podolski y Rosen, de Bohr, y de Schrodinger de 1935 están basadas en experimentos pensados, y se acompañan de preguntas como: ¿Son la posición y la cantidad de movimiento parte de la realidad física? ¿La medición de una propiedad de una partícula determina la medición de la propiedad de otra partícula con la que está correlacionada, e involucra

senhales mas rapidas que la de la luz? Existen estados enredados superposiciones de gato vivo y nucleo de uranio, y de gato muerto y nucleo de torio mas particula alfa? Se pueden realizar experimentos con un solo atomo? Se pueden hacer experimentos con un solo foton? Existen los saltos cuanticos? Se puede observar el colapso de la funcion de onda? En contraste, las otras tres conferencias involucran experimentos planeados y realizados con objetivos muy concretos que dan respuestas a tales preguntas: estados fotonicos enredados a distancias cada vez mayores; experimentos complementarios con un solo ion cuyos grados de libertad de posicion en una trampa y de excitacion interna estan enredados, y estados de un solo foton en una cavidad de resonancia en interaccion con atomos de Rydberg enredados; saltos cuanticos en atomos de Rydberg en campo electrico, colapso de la funcion de onda, coherencia y decoherencia.

Como resumen de esta seccion podemos afirmar que en los parrafos primero, tercero y cuarto se han dado las descripciones de las cuatro conferencias de los tercios sucesivos del ciclo. El segundo parrafo consiste en una explicacion general de los papeles de los experimentos y de la formulacion matematica de las leyes en la fisica, y de sus relaciones en el avance de cada area. Las conexiones del primer tercio y los otros dos se pueden entender como la aplicacion de los fundamentos que han permitido predecir y corroborar la existencia de nuevas formas de luz, materia e interacciones, en momentos diferentes, y en escalas macroscopicas y elementales, respectivamente. Los ultimos a su vez refuerzan la validez de los fundamentos comunes.

III Fundamentos Comunes de Electromagnetismo y de Mecanica Cuantica

En esta seccion discutimos con cierto detalle el Principio de Superposicion, el Problema de Eigenvalores, el uso de los Potenciales Electromagneticos, y la Estructura Lineal y Cuadratica como elementos comunes en Electromagnetismo y Mecanica Cuantica. La discusion esta orientada a reconocer las conexiones entre los elementos en una y otra area, como se presento en las conferencias sucesivas.

Las ecuaciones de Maxwell describen la radiacion electromagnetica, incluida la luz, como un fenomeno ondulatorio con densidades de energia y de cantidad de movimiento distribuidas en el espacio y el tiempo (1865). El electromagnetismo fue el precursor de la mecanica cuantica (1900,1905) y de la teoria de la relatividad (1905). La ecuacion de onda de Schroedinger describe las propiedades ondulatorias de la materia, propuestas en la hipotesis de de Broglie (1923), y explica la cuantizacion como un problema de eigenvalores (1926). Ambos tipos de ecuaciones comparten las propiedades de ser lineales en las derivadas espaciales y temporal, y en consecuencia sus respectivas soluciones satisfacen el principio de superposicion. A continuacion se discuten los cuatro temas descritos en el parrafo anterior, proporcionando respuestas detalladas a las preguntas respectivas formuladas en la Seccion II.

Las ecuaciones de Maxwell son la expresion matematica de las leyes de la electrodinamica: La ley de Gauss electrica reconoce que las cargas electricas son fuentes de flujo electrico. La ley de Gauss magnetica reconoce que no hay fuentes de flujo magnetico. La ley de Faraday reconoce que la rapidez de cambio con respecto al tiempo del flujo magnetico es fuente de circulacion electrica. La ley de Ampere-Maxwell reconoce que hay dos fuentes de circulacion magnetica: corrientes de cargas electricas y la rapidez de cambio con respecto al tiempo del flujo electrico. Adicionalmente, se tiene la conservacion de carga, bajo la forma de la ecuacion de continuidad: divergencia de la densidad de corriente mas la derivada parcial con respecto al tiempo de la densidad de carga es cero. Para cada distribucion de cargas y corrientes, los campos electricos y magneticos son soluciones de las ecuaciones de Maxwell. El principio de superposicion establece que para una superposicion de diversas

distribuciones de cargas y corrientes, los campos respectivos son la misma superposicion de los campos individuales.

La ecuacion de Schroedinger describe la dinamica cuantica de sistemas materiales, igualando la accion del operador Hamiltoniano con la accion de la derivada parcial con respecto al tiempo sobre la funcion de onda. El Hamiltoniano tiene un termino de energia cinetica, proporcional al operador Laplaciano, y un termino de energia potencial. Para cada potencial, la ecuacion tiene una base completa de soluciones ortonormales. El principio de superposicion establece que cualquier superposicion de las soluciones de la base es tambien una solucion. Cada funcion de onda es una amplitud de probabilidad, el producto con su compleja conjugada es una densidad volumetrica de probabilidad, y su integral volumetrica es cien por ciento o uno.

Para identificar la conexion entre los principios de superposicion en los temas respectivos regresamos al caso del electromagnetismo. Efectivamente, los sistemas electromagneticos incluida la radiacion electromagnetica tienen una densidad de energia distribuida en el espacio, proporcional a la suma de los productos escalares de los campos electricos y magneticos con sus respectivos complejos conjugados. La integral volumetrica de esa densidad es la energia total del sistema electromagnetico; en particular para la radiacion electromagnetica de una frecuencia dada esa energia es nhf , un multiplo entero del cuanto de luz, segun Einstein (1905). La conexion buscada se reconoce al comparar las integrales de los dos ultimos parrafos, y tomando la de la radiacion como energia promedio por foton para que sea uno. Esta conexion es tambien consistente con la posibilidad de escribir las ecuaciones de Maxwell como una ecuacion de tipo Schroedinger o de Dirac con un Hamiltoniano lineal en el momento lineal contraida con una matriz de espin uno, y una funcion de onda que es el bivector $E + iB$, segun Silverman (1908).

Tanto las ecuaciones de Maxwell como la de Schroedinger son separables en el sentido de que sus soluciones son factorizables en sus partes espaciales y temporal, y las ecuaciones mismas se transforman en ecuaciones diferenciales ordinarias de problemas de eigenvalores. Los eigenvalores pueden estar cuantizados segun las condiciones de frontera. Esto es bien conocido en la mecanica cuantica, y aqui nuestro objetivo es ilustrar su ocurrencia en electromagnetismo. Efectivamente, las ecuaciones de Maxwell son ecuaciones de primer orden en las derivadas espaciales y temporal, y acoplan los campos electricos y magneticos en las leyes de Faraday y Maxwell. Su desacoplamiento se logra tomando el rotacional de ambas ecuaciones, y usando sus companheras para llegar a la ecuacion de onda clasica para ambos campos, con fuentes que son el gradiente de la densidad de carga y la derivada parcial de la densidad de corriente, y el rotacional de la densidad de corriente, respectivamente. Las ondas electromagneticas son soluciones de estas ecuaciones en regiones donde no hay fuentes y las ecuaciones son homogeneas. Para fuentes con variaciones armonicas en el tiempo las soluciones son bases de Fourier y la parte espacial obedece la ecuacion de Helmholtz. Si las fuentes se localizan en una superficie cerrada, paralelepipedo rectangular, cilindro, esfera, etc. la ecuacion se puede resolver por el metodo de separacion de variables en los respectivos sistemas de coordenadas, a partir de las soluciones de la ecuacion de Helmholtz bien comportadas en el interior y el exterior del volumen limitado por la superficie, sujetas a las condiciones de frontera del sistema fisico por analizar y a las leyes de Maxwell en su forma de condiciones de frontera en la superficie. En el caso de cavidades de resonancia la condicion de frontera es la anulacion de los campos en el exterior, y en el caso de antenas la condicion asintotica requiere ondas salientes. El confinamiento de las fuentes en la superficie se traduce en la cuantizacion de las frecuencias y tambien de las energias de las radiaciones estacionarias y viajantes en ambos dispositivos. El efecto es matematicamente equivalente a lo que ocurre en sistemas materiales.

Aharonov y Bohm publicaron el trabajo “Significado cuantico de los potenciales electromagneticos” en 1959, prediciendo que el patron de interferencia de dos haces de electrones que se mueven alrededor de un solenoide recto con corriente en circulos paralelos y campo uniforme en su interior cambia periodicamente con el valor del flujo magnetico, dependiente de la circulacion del potencial vectorial magnetico alrededor del solenoide. En 1961, Chalmers implemento experimentalmente la observacion de este efecto Aharonov-Bohm.

En la conferencia del Ciclo, el Autor decidio incluir tambien el significado clasico del potencial vectorial, tomando en cuenta que es un concepto que no es bien entendido ni explicado en los libros de texto. Su origen y nombres han sido manejados por fisicos reconocidos: Faraday lo concibio como el estado electrotonico, Maxwell lo identifico y llamo potencial de momento electrocinetico, y Feynmann y Konopinski argumentan que es un potencial de cantidad de movimiento por unidad de carga en la interaccion de una carga electrica en presencia del mismo, y tambien un potencial de energia por unidad de carga y unidad de velocidad paralela al potencial. En los libros la ultima frase se caracteriza como la “prescripcion de acoplamiento minimo”.

Esta version escrita de la conferencia esta enfocada en explicar el significado fisico clasico del potencial vectorial, su conexion con el significado cuantico reconocido por Aharonov y Bohm, y en la discusion sobre cual de las descripciones del electromagnetismo, en base de campos de fuerza y de potenciales electromagneticos, es mas fundamental.

Efectivamente, si tomamos la fuerza de Lorentz como el punto de partida reconocemos su componente electrica, producto de la carga y el campo de intensidad electrica, y su componente magnetica producto de la carga dividida entre la velocidad de la luz y el producto vectorial de su velocidad y el campo de induccion magnetica. Por otra parte el campo de induccion magnetica, por ser solenoidal segun la ley de Gauss, se puede escribir como el rotacional del potencial vectorial electromagnetico; y a su vez por la ley de Faraday, el campo de intensidad electrica se puede escribir como la superposicion de un gradiente de potencial escalar y de la derivada parcial con respecto al tiempo del potencial vectorial. La sustitucion de estas conexiones entre campos de fuerza y de potenciales en la fuerza de Lorentz, permite reescribirla como la superposicion de un termino de gradiente negativo y un termino de derivada total negativa con respecto al tiempo, como el lector interesado puede hacerlo. Desde luego el termino de la fuerza magnetica, incluye el triple producto vectorial de la velocidad y el rotacional del potencial vectorial, cuyo desarrollo incluye un termino de gradiente de la contraccion de la velocidad y el potencial, y el de la contraccion de la velocidad con el gradiente del potencial; el ultimo termino combinado con la derivada parcial temporal del potencial asociado a la fuerza electrica da la derivada total temporal correspondiente. El resultado neto conecta con la interpretacion de los potenciales de energia y de cantidad de movimiento asociada a Feynmann y Konopinski en el penultimo parrafo. Y le da una base fisica y matematica a la “prescripcion”

Una vez que entendemos el significado fisico del potencial vectorial en el nivel clasico, podemos analizar su conexion con su significado cuantico. En la dinamica Hamiltoniana, cantidad de movimiento y coordenada de posicion son cantidades canonicamente conjugadas, y lo mismo ocurre con energia y tiempo; sus productos respectivos tienen las mismas dimensiones de accion. La constante de Planck es la unidad natural de accion, y en el nivel cuantico la accion esta cuantizada. La descripcion cuantica ondulatoria de los electrones se puede hacer en base de una onda plana representada como una exponencial imaginaria y una fase que es la diferencia de las dos acciones de la frase anterior dividida entre la constante de Planck entre 2π , quedando medida en radianes. Su parte

espacial depende del número de onda, según de Broglie, y su parte temporal depende de la frecuencia según Schrödinger.

El efecto Aharonov-Bohm involucra la acción asociada a la interacción del electrón con el potencial vectorial externo al solenoide, donde el campo de fuerza magnética es nulo. El patrón de interferencia depende de la diferencia de fases entre los haces que viajan de un lado y otro del solenoide, que resulta equivalente a la integral de circulación alrededor del solenoide al cambiar la dirección de la integración y el orden de los límites en el término negativo. Por el teorema de Stokes, esa circulación es la integral de área en la sección del solenoide del rotacional del potencial, que a su vez es la misma integral del campo magnético, o sea el flujo magnético en el interior del cilindro. Esta situación también ha dado lugar al debate sobre cuáles son los conceptos fundamentales: campos de fuerza o campos de potencial. Y también sobre no localidad, y lo que sucede bajo transformaciones de norma.

Para dar perspectiva al debate es importante recordar que la dinámica clásica se desarrolló en las modalidades Newtoniana en base a fuerzas y trayectorias en el espacio, Lagrangiana en base de energía potencial y espacio de configuración de coordenadas y velocidades generalizadas, y Hamiltoniana en base de energía potencial y espacio fase de coordenadas generalizadas y cantidades de movimiento canónicamente conjugadas. Sus formalismos matemáticos son diferentes pero su contenido físico es común, y sus ecuaciones de movimiento son equivalentes. Las conexiones entre fuerzas y energías potenciales también están perfectamente entendidas. Estas ideas se extienden a la dinámica electromagnética, y también en la mecánica cuántica y las teorías de campo cuánticas, como se discutió en el Ciclo de conferencias sobre Transformaciones de Norma y el Principio de Invariancia de Norma. Las formulaciones alternativas no son excluyentes sino complementarias, y lo más importante es reconocer sus conexiones.

El ensayo de Dyson de la cuarta conferencia consta de dos partes: La Modestia a veces no es una Virtud, y la Teoría Electromagnética y la Mecánica Cuántica. Su propósito está enfocado en reconocer las dificultades en entender cada una de ellas, y que el entendimiento de la primera puede ayudar al entendimiento de la segunda. A continuación se describen las ideas básicas de una y otra, acompañadas de comentarios del Autor.

Dyson identifica La Evolución de las Especies de Darwin como el trabajo científico más importante del siglo XIX, y la Teoría Electromagnética de Maxwell como el trabajo más importante de la Física. También considera que Maxwell en su conferencia en la Royal Society, Sección de Física y Matemáticas, debió seguir el ejemplo de Newton dos siglos antes, y decir las ecuaciones de Maxwell rigen el comportamiento de la radiación electromagnética en el universo. En vez de eso habló de vórtices electromagnéticos, cuyas matemáticas no fue entendida por los físicos y cuya física no era entendida por los matemáticos.

Para entender la selección de Maxwell del tema de su presentación es importante leer el prólogo en su Tratado de Electromagnetismo, y hablar de uno de los problemas que era de mayor interés en esa época. En el prólogo, Maxwell admite que él no ha realizado experimentos de electricidad y magnetismo, y afirma que su propósito es expresar los resultados de Faraday en su propia obra Investigaciones Eléctricas Experimentales, en el lenguaje de las matemáticas que se usan para describir los fenómenos físicos. El problema de interés en la Óptica y por extensión en el Electromagnetismo era entender el medio en que se propaga la luz y la radiación electromagnética, que desde cincuenta años antes se trataba de entender en base de la hipótesis del éter luminífero. Las investigaciones de Maxwell incluyeron construir un modelo mecánico de tal éter, que no fructificaron. Helmholtz fue quien estuvo en posición de entender el trabajo de Maxwell al afirmar: la teoría electromagnética está en las ecuaciones de Maxwell, y sus soluciones para la radiación electromagnética son vórtices ideales sin

efectos de disipación. Hertz, alumno de Helmholtz, fue quien produjo experimentalmente ondas electromagnéticas en su laboratorio en 1887, corroborando las predicciones de Maxwell.

En la segunda parte del ensayo, Dyson identifica el nivel de las soluciones de las ecuaciones de Maxwell en términos lineales de los campos de intensidad eléctrica y de inducción magnética, en contraste con el nivel de sus combinaciones cuadráticas asociadas a sus propiedades dinámicas de densidad de energía y densidad de cantidad de movimiento. Su argumentación está basada en consideraciones dimensionales para darle preferencia al segundo nivel. Concretamente, la densidad de energía se mide en joules entre metro cúbico, y si queremos conectar con el nivel lineal es necesario extraer una raíz cuadrada con unidades de raíz cuadrada de joule entre raíz cuadrada de metro cúbico; y que significado tienen tales raíces cuadradas? Su conclusión es que el segundo nivel es observable o medible, y el primero no.

La comparación con el caso cuántico reconoce también los dos niveles: lineal en la función de onda o amplitud de probabilidad, y cuadrático en la densidad de probabilidad, en valores esperados, y en elementos de matriz. Ahora la unidad de la densidad de probabilidad es uno entre metro cúbico, y la de su raíz cuadrada para la amplitud de probabilidad es uno entre raíz cuadrada de metro cúbico.

Conclusión: la densidad de probabilidad es observable o medible, y la amplitud de probabilidad no.

El comentario del Autor es válido para el electromagnetismo y la mecánica cuántica, afirmando que es en el primer nivel que la evidencia del comportamiento ondulatorio de la luz fue observada en su difracción por Francisco Grimaldi en la primera mitad de los 1600s, Young, Fresnel y Fraunhofer a principios de los 1800s; y de los electrones en los experimentos de difracción de G.P. Thompson, y de Davisson y Germer, en 1927 validando la hipótesis ondulatoria de la materia de de Broglie.

Adicionalmente, la unidad del campo de intensidad eléctrica es dina entre unidad electrostática y la unidad del campo de inducción magnética es dina entre unidad de carga electromagnética y centímetro entre segundo, equivalentes a las mencionadas en el penúltimo párrafo.

El Autor coincide con Dyson en revisar conjuntamente y comparativamente las dos disciplinas para apreciar sus fundamentos comunes, y entender mejor ambas.

Percolación y sus variaciones: transiciones geométricas

F. Leyvraz

*Instituto de Ciencias Físicas—Universidad Nacional Autónoma de México, Cuernavaca, Morelos, México
and Centro Internacional de Ciencias, Cuernavaca, Morelos México**

(Dated: 7 de octubre de 2021)

Aquí damos una presentación somera de algunos aspectos de la teoría de percolación, y del tema relacionado de la teoría de grafos, también conocidos como *redes*. Tocamos los temas siguientes: la percolación, dos modelos de formación de redes—el llamado modelo de crecimiento preferencial de Barabási y Albert por un lado, y el modelo puramente aleatorio de Erdős y Rényi del otro. Discutiremos la estrecha relación ente éste último y la percolación: el modelo de Erdős–Rényi resulta ser la teoría de campo medio de la percolación. Ilustraremos algunos conceptos importantes de percolación usando la teoría de campo medio. Finalmente discutiremos un modelo extraño llamado *percolación explosiva*, que por un tiempo aparentó tener una transición de percolación discontinua.

I. GEOMETRÍA Y CONEXIONES EN FÍSICA Y OTRAS CIENCIAS

En lo siguiente quiero fijarme antes de todo en las *conexiones* entre partículas (u otros objetos). ¿Dónde viene el interés en esto? Más adelante veremos que el concepto de conexión juega un papel muy importante en muchos contextos. Pero antes de todo, quiero mostrar que, aún en os sistemas más sencillos, fijarse en conexiones da lugar a una complejidad inesperada.

Una figura dice más que mil palabras: en las siguientes muestro 3 redes ocupadas *al azar* por puntos, con densidades muy parecidas. No se ve nada de interés. Por otro lado, si busco los agregados *conexos*, es decir aquéllos en los que cada partícula está conectada con cada otra, vemos algo que, a primera vista parece sorprendente (véase la Fig. 1). Vemos un cambio muy rápido en la naturaleza y el tamaño de este agregado: inicialmente, en la primera figura, está muy ramificado, casi no incluye “lagos” grandes de lugares vacíos. Después, en la segunda, se torna más denso, pero todavía contiene sólo un porcentaje ínfimo de los puntos. Ahora sí aparecen lagos de variados tamaños, y una estructura irregular, similar a todas escalas, que se llama *fractal*. Los fractales aparecen en muchos sistemas naturales y han sido objeto de un sinnúmero de estudios. Un libro magnífico al respecto el del matemático francés Mandelbrot [1]. Otro, que muestra técnicas específicas para crear fractales, se debe a Barnsley [2] y muestra la inmensa riqueza de la geometría fractal y como los podemos ver en un sinnúmero de sistemas naturales. Finalmente, en la tercera figura se torna denso, con muchos lagos, pero todos de un tamaño relativamente menor. Esto es también el primer caso en que el número de partículas contenidas en el agregado mayor crece proporcionalmente al número de partículas en el sistema. En otras palabras, una fracción finita de los sitios ahora aparecen como parte del agregado infinito.

Vemos entonces un *cambio repentino* cuando nos fijamos en la geometría de algo tan sencillo como un conjunto de puntos sin correlaciones (gas ideal). Tenemos una *transición de fase*. Y más específicamente, se trata de una transición de fase de *segundo orden*, en la que la naturaleza de un sistema cambia profundamente, pero de manera continua: pasamos de una fase desconectada a bajas densidades, a otra que es plenamente conectada a densidades más altas, y pasando por una fase intermedia caracterizada por la existencia de estructuras a cada escala, estructuras fractales y auto-similaridad.

Ahora ¿por qué me ha de interesar? ¿Sirve esto en el mundo real?

Una aplicación importante es sin duda el estudio de medios porosos: medios que están zurcados por vacíos abundan en muchos contextos. Rocas, arenas, polvos, pero también geles con alta porosidad para obtener una alta resistencia mecánica con poco peso (aerogeles) o en catalizadores para poder realizar con eficiencia una reacción química que sólo se puede dar a la superficie del catalizador, por lo que requerimos de una cantidad de poros tan grande como sea posible. Otro ejemplo es la conductividad eléctrica en una mezcla de metal con aislante: en estos casos, el flujo sólo puede suceder a través de las conexiones. En todos estos casos, el estudio de la geometría y de las conexiones es fundamental. En transiciones de fase, como son la transición líquido–gas, también el estudio geométrico puede arrojar mucha luz sobre lo que pasa: se identifican gotas de ambas fases, y se puede estudiar como cambian sus tamaños y la variación de sus propiedades [3] conforme pasamos por un cambio de fase físico.

Pero en otras disciplinas abundan los sistemas donde el estudio de conexiones es central. Investigar las redes que resultan de un sistema de conexiones puede hacer en muchos contextos. Demos algunos ejemplos:

*Electronic address: leyvraz@icf.unam.mx

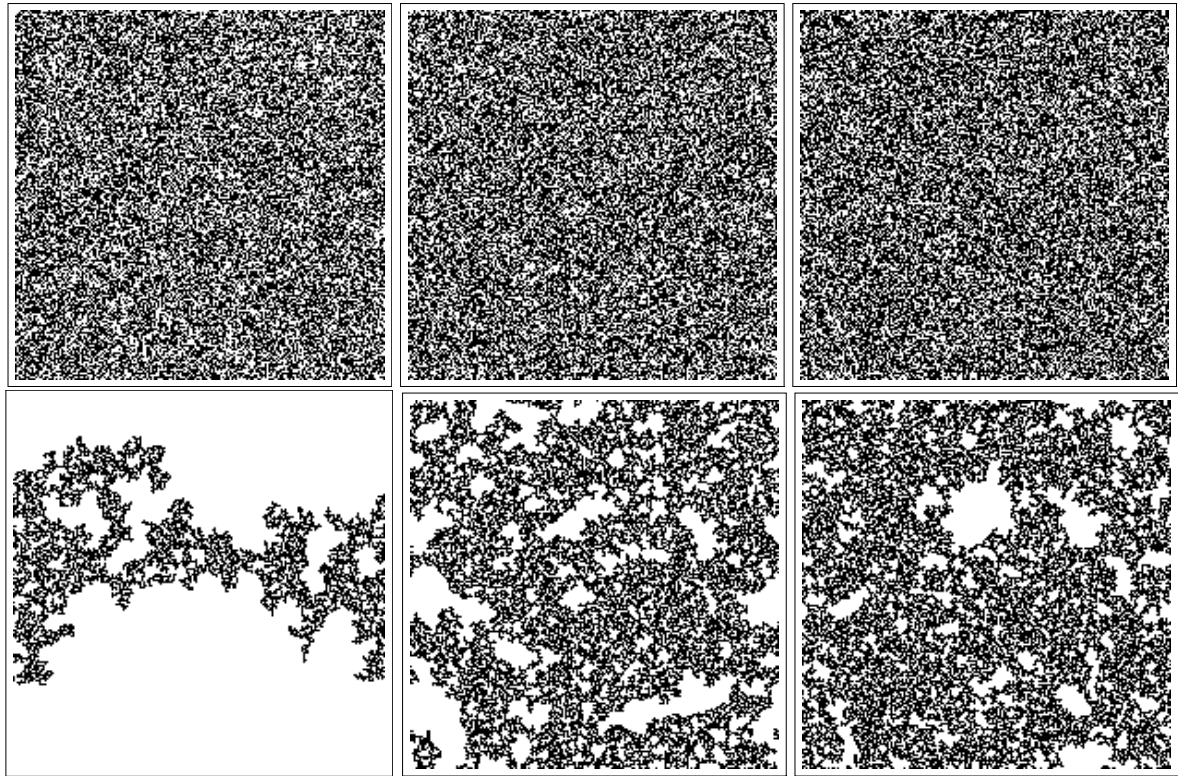


Figura 1: Arriba muestro 3 redes ocupadas al azar con densidades de partículas 0.58, 0.593 y 0.61. Abajo muestro el *mayor agregado conectado* en cada una de las 3 redes. En otras palabras, busqué el conjunto de partículas más grande con la propiedad que todas las partículas estén conectadas entre sí por un camino.

- Epidemiología: los elementos son las personas, y el vínculo entre ellas es el contagio.
- Análisis del lenguaje: los elementos son las palabras de un texto, y se conectan dos palabras que aparecen juntas o cerca.
- Las redes sociales y el Internet proveen directamente ejemplos obvios de redes.
- Señales temporales: se conectan si la correlación entre 2 señales es mayor que un umbral dado.
- ...

Ahora podemos preguntar: claro, se puede añadir una estructura de red a cualquier sistema. Pero ¿a qué nos lleva esto? Resulta que para cualquier red podemos definir varias cantidades de interés, de modo que, definida una red, se sugiere de inmediato el análisis de varios problemas. Podemos, por ejemplo, estudiar como varía el número de vecinos de cada nodo, mirando la *distribución* de este número. Este número ¿tiene más o menos siempre la misma magnitud? ¿O al contrario existen *hubs*—nodos focales—con un número muy grande de vecinos, alrededor de los cuales se organiza la red? O podemos ver cuántas partes conexas hay, y cuál es la distribución de sus tamaños. Por otro lado, se puede investigar en cuánto cada nodo contribuye a la conexión de la red. Existen muchas manera de medir esto, siendo la más conocida el llamando algoritmo de PageRank de Google. Finalmente, podemos mirar cómo es la topología de las conexiones. ¿Forma la red un árbol, sin ciclos? O al contrario ¿hay ciclos pequeños como por ejemplo triángulos, permitiendo así la identificación de grupos altamente conectados en la red? Otro efecto curioso que se observa en varias redes es el de *mundo pequeño*: resulta que en muchos casos el camino más corto para conectar dos puntos arbitrarios de la red es bastante corto, específicamente, crece como $\ln N$ en función del número N de nodos en la red. Este efecto se dio primero a conocer por la investigaciones de Milgram [4], quien encargó a unos sujetos en Estados Unidos enviar una carta a otra persona escogida al azar en Estados Unidos, mandándole una carta a una persona conocida encargándole reenviar la carta a otra persona conocida, intentando acercarse a la meta. Resultó que en promedio no se requirieron más de 6 pasos para lograr la conexión. Este extraño hallazgo experimental se puede descubrir como una propiedad de muchos modelos matemáticos de redes.

Por motivos obvios nos tendremos que limitar en los temas que tratamos. En la Sección II presentamos dos modelos de características, que ambos generan unas redes interesantes, y describiremos sus propiedades de manera algo escueta. El uno es el modelo de Barabási–Albert [5, 6], también conocido como modelo de crecimiento preferencial. El otro es un modelo de red formada enteramente al azar. Este último modelo se conoce como el modelo de Erdős–Rényi. En la Sección III discutiremos el papel que éste último juega en el estudio de la percolación, que corresponde al modelo que vimos al principio del artículo: veremos que se trata de lo que llamamos la teoría de campo medio de la percolación. En la sección IV hablaremos de un modelo desarrollado más recientemente [7], llamado percolación explosiva. Este modelo también genera una red en algunos aspectos similar a la red de Erdős–Rényi, pero con la particularidad que la transición, en vez de ser continua, como lo es la transición de percolación, es tan brusca que dicha continuidad no se puede verificar numéricamente. En la sección V presentaremos unas conclusiones.

II. MODELOS DE REDES

Como vimos, es posible generar una gran cantidad de redes con datos reales. Pero ¿hay modelos teóricos con los que se pueden comparar estos datos? En realidad hay muchísimos, y nos limitaremos a dos. El primero es un modelo debido a Barabási y Albert, que obtiene una estructura asombrosamente compleja con unas reglas de extrema sencillez. El otro es el modelo más elemental de una red perfectamente aleatoria. El interés de este modelo, como veremos, es que representa una aproximación muy interesante a la percolación común y corriente que describimos al principio en la Figura 1. En la Sección IV discutiré otro modelo con propiedades algo extrañas: la percolación explosiva.

A. Crecimiento por vinculación preferencial

El modelo de Barabási–Albert, o modelo de crecimiento preferencial, se describe del modo siguiente. La red crece en el tiempo: se le añade a cada paso un punto nuevo, empezando típicamente con uno solo, usando las reglas siguientes:

1. Se elige un punto P al azar de la red ya formada.
2. Con una probabilidad dada p se vincula el punto nuevo al punto P .
3. Con probabilidad $1 - p$ se elige al azar un punto Q conectado con P y se conecta el nuevo punto a Q .

La probabilidad p es un parámetro del modelo que se puede fijar arbitrariamente.

A primera vista, puede parecer que los dos casos, conectar el punto nuevo a un punto arbitrario según la regla 2, o conectarlo a un vecino de un punto arbitrario, no tienen mayor diferencia. Veamos que sí hay una diferencia notable: claramente, elegir un punto al azar no tiene sesgo alguno y todos los puntos aparecen con igual probabilidad. Por otro lado, si elegimos un vecino al azar de un punto arbitrario, entonces se elegirán con mayor frecuencia a los puntos que ya tienen muchas conexiones. Este resultado, a primera vista algo contraintuitivo, se ilustra gráficamente en la Figura 2.

El modelo, por lo tanto, conlleva una *inestabilidad*: si, durante el crecimiento de la red, un nodo tiene, por casualidad, más conexiones que otro, se elegirá con mayor frecuencia para que se le conecte un punto nuevo, por lo que su ventaja se acrecenta (esto a veces se llama el *efecto Mateo*, haciendo referencia a Mateo 13:12: “Porque a cualquiera que tiene, se le dará, y tendrá más; pero al que no tiene, aun lo que tiene le será quitado.”). Sin embargo, resulta que la ventaja así adquirida no es tan fuerte como para que un solo nodo elimine a la competencia de manera definitiva. De este modo aparecen nodos con número de vecinos altamente variables. Este número no tiene un valor característico, sino que varía como una *ley de potencia*.

En fórmulas: la probabilidad que un nodo tenga k vecinos decae como

$$\text{Prob}(\text{número de vecinos} \geq k) \simeq \text{const.} \cdot k^{-\alpha}. \quad (1)$$

Resulta que α depende continuamente de p y tiene un valor arriba de 3 [8]. Esto significa que existen *hubs*, es decir nodos con muchos vecinos, que ayudan a conectar muchos nodos con menos conexiones entre sí.

En datos reales esto se observa en muchos sistemas. La red de aeropuertos es un ejemplo donde la palabra *hub*—en el sentido de aeropuerto central—apareció posiblemente antes del concepto correspondiente en la teoría de redes. Otros ejemplos son: la red describiendo las conexiones del internet, o de otras redes sociales, así como las redes de citas o de coautoría en artículos científicos. Un ejemplo de la concentración de citas a artículos se ve en el hecho siguiente: cada artículo que se publica cita—a lo menos—10 artículos. Sin embargo, aún artículos en revistas prestigeadas no reciben más que 2 citas en promedio. Para que esto funcione, es claramente necesario que la gran mayoría de los artículos citados se concentre en unos pocos. Si quieres ver mayores detalles al respecto, hay detalles interesantes en [9].

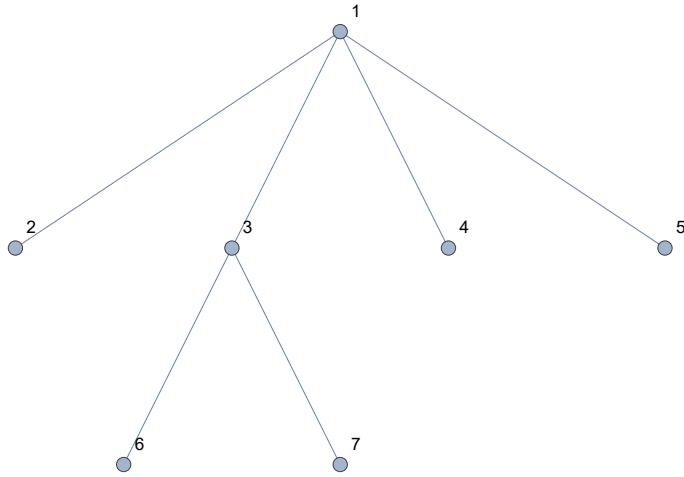


Figura 2: Al elegir un punto al azar, todos los puntos, de 1 a 7, tienen la misma probabilidad de ser elegidos. Pero si se elige un punto al azar, y luego se busca a uno de sus vecinos, es claramente más probable que este vecino resulte ser 1 que, por ejemplo, 5. Lo primero tiene probabilidad $10/21=0.48$, lo segundo $1/28=0.036$.

Hagamos una corta reseña de las otras propiedades de las que hablábamos inicialmente: la red que resulta de esta construcción es siempre, por definición, un árbol, de manera que no hay ciclos, ni pequeños ni grandes. En cuanto al efecto de mundo pequeño, se prueba que en una red de tipo Barabási–Albert con N elementos, siempre se pueden conectar dos puntos arbitrarios con un camino de $\ln N / \ln \ln N$ pasos [10], es decir, el efecto de mundo pequeño es aún algo más fuerte que en la red de Erdős–Rényi, donde la longitud del camino mínimo es de $\ln N$.

B. Red de Erdős–Rényi

Recordemos el modelo que habíamos visto al principio de este artículo: en una red cuadrada habíamos ocupado al azar con probabilidad p cada nodo de la red, y habíamos considerado los agregados *conectados*. En otras palabras, habíamos conectado los puntos ocupados vecinos, e investigamos luego los tamaños de las componentes conexas de la red que resulta.

En lo siguiente trataremos el modelo de Erdős–Rényi originalmente definido en [11]. Éste es el modelo más sencillo de una red que es solamente azarosa, en el sentido que no tiene estructura alguna. La construcción es sencilla: se toman un gran número de puntos, digamos N . Cada par de puntos se conecta con probabilidad p y se deja sin conectar con probabilidad $1 - p$, independientemente de lo que suceda con los demás pares de partículas.

El número promedio de vecinos en este caso es de Np . Si tomamos un valor fijo de p , esto lleva a una situación donde el número promedio de vecinos tiende a infinito para redes grandes. Esto difiere notablemente de las redes típicas en la realidad. Busquemos entonces hacer que el número promedio de vecinos se mantenga constante, digamos igual a z , para cada N , y por ello fijamos

$$p = \frac{z}{N} \tag{2}$$

En este modelo, se prueba que la distribución del número de vecinos de un nodo es Poissoniana, es decir, en fórmulas

$$\text{Probabilidad que el número de vecinos sea igual a } k = \frac{N!}{k!(N-k)!} \left(1 - \frac{z}{N}\right)^{N-k} \left(\frac{z}{N}\right)^k \approx \frac{z^k}{k!} e^{-z}, \tag{3}$$

donde la última igualdad aproximada es válida al tender $N \rightarrow \infty$. Esto nos dice que cada nodo tiene aproximadamente z vecinos, y que sólo hay poca variación: la variancia, por ejemplo, también es igual a z . En otras palabras no hay *hubs*.

Pero sí hay una transición repentina: para cada z el sistema consta de varias componentes conexas (existe, como lo mostraron en sus trabajos Erdős y Rényi, un p tal que con probabilidad uno el agregado esté totalmente conectado. Sin embargo, esto sucede sólo cuando p crece más rápido que $\ln N/N$, es decir, en un caso donde el número promedio de vecinos de un nodo diverge cuando $N \rightarrow \infty$, que es un caso que no consideramos). La transición de la que hablo es la

siguiente: si $z < 1$, entonces la red consta de componentes conexas cuyo tamaño no crece con N . Ningún componente tiene un tamaño que diverja con N .

Por otro lado, si $z = 1$, entonces el tamaño del mayor agregado crece con probabilidad 1 con $N^{2/3}$, y si $z > 1$, crece proporcionalmente a N . En otras palabras, una proporción finita del sistema está conectado entre sí cuando $z > 1$.

El modelo de crecimiento preferencial se dio en términos de un mecanismo *dinámico* de crecimiento, es decir, describimos la manera en que el sistema crecía en función del tiempo. En el caso de la red de Erdős–Rényi podemos hacer lo propio. Esta vez, sin embargo, empezamos con N puntos, no con uno solo. Las reglas son entonces de gran sencillez: en cada paso hacemos las dos acciones siguientes

1. escoger 2 puntos al azar y conectarlos
2. incrementar el tiempo en $1/N$

El sistema a tiempo t corresponde entonces, como es fácil darse cuenta, a una red de Erdős–Rényi con $z = t$. Ilustramos lo que pasa con la red de Erdős–Rényi con z cercana a 1 en la Figura 3: en la figura de arriba mostramos un caso con z algo menor que 1, y abajo con z algo mayor que 1. $N = 1000$ en ambos casos. Observamos un cambio notable en el tamaño del mayor agregado, que pasa de tener tamaño 78 a 202 con un cambio menor en z . Similarmente se nota una baja en el número de agregados pequeños, y un aumento en la brecha que separa el agregado mayor del siguiente: las indicaciones de una transición repentina que los trabajos de Erdős–Rényi [11] describen de manera matemáticamente rigurosa.

III. LA RED DE ERDÖS–RÉNYI Y LA PERCOLACIÓN

Aquí quiero recalcar la similitud entre las redes de Erdős–Rényi y la percolación. Recordemos: la percolación consiste en poner una fracción p de nodos ocupados en una red regular: cuadrada, cúbica u otra e identificamos los agregados conectados. El problema del estudio de estos agregados es de gran dificultad, y ha dado lugar a una gran variedad de estudios tanto matemáticos como numéricos. No puede ser nuestro propósito dar cuenta de ellos, pero sí indicar el libro de Stauffer [12] como un muy buen punto de partida para averiguar algo al respecto. Aquí averiguamos cuántos hay, y de qué tamaños son.

En este contexto podemos definir $n_s(p)$ como el número de agregados de tamaño s dividido por el número total de partículas. En una red cuadrada por ejemplo

$$n_1(p) = p(1-p)^4 \quad (4)$$

$$n_2(p) = 4p^2(1-p)^6 \quad (5)$$

Esto se explica del modo siguiente: para tener un agregado de un solo punto, se requiere que el punto esté ocupado, lo que tiene probabilidad p , y que cada uno de sus 4 vecinos esté vacío, lo que en total tiene probabilidad $(1-p)^4$.

Estos números $n_s(p)$ se pueden en principio calcular para cada valor finito de s , pero las expresiones se vuelven rápidamente inmanejables: son polinomios de orden muy alto, con coeficientes racionales con numeradores y denominadores inmensos. Este enfoque no resulta útil para obtener resultados sencillos, aunque sí se usa como una técnica numérica.

La percolación es un tema complejo, y no hay soluciones exactas, pero sí hay muchos estudios numéricos. Es similar, por ejemplo, al modelo de Ising para el ferromagnetismo, con la diferencia que para el modelo de Ising sí existe una solución exacta no-trivial: el modelo de Ising en dos dimensiones.

Para el modelo de Ising, hay una *aproximación* muy útil, que es la de Curie–Weiss, o la teoría de campo medio. La aproximación similar para la percolación resulta ser una descripción exacta para las redes de Erdős–Rényi.

En efecto, vemos que la única diferencia entre la percolación y las redes de Erdős–Rényi es que las conexiones en percolación sólo ocurren entre vecinos de una red, mientras no existen tales limitaciones en el modelo de Erdős–Rényi. De manera similar, para el modelo de Ising, las interacciones tienen lugar entre vecinos cercanos, mientras en la teoría correspondiente de campo medio, cada partícula interactúa (débilmente) con cualquier otra.

Para identificar el agregado al que pertenece un determinado punto, identificamos todos sus vecinos, y después los vecinos de sus vecinos, y así sucesivamente de generación a generación. Ya que el número de vecinos es de z en promedio y que $z \ll N$, la probabilidad que en el proceso de identificar el agregado, tengamos un ciclo, es despreciable, a lo menos si el agregado es finito. La figura 4 ilustra la idea: empezamos con un nodo arbitrario, la raíz, y determinamos los descendientes.

Esto permite analizar el problema usando *recurrencias*. Por ejemplo, consideremos la probabilidad de conectarse al infinito. Intuitivamente, si el número promedio de conexiones z es menor que 1, al identificar el agregado de la manera arriba descrita, vemos que el número promedio de descendientes de generación k de un nodo dado es de z^k , lo cual

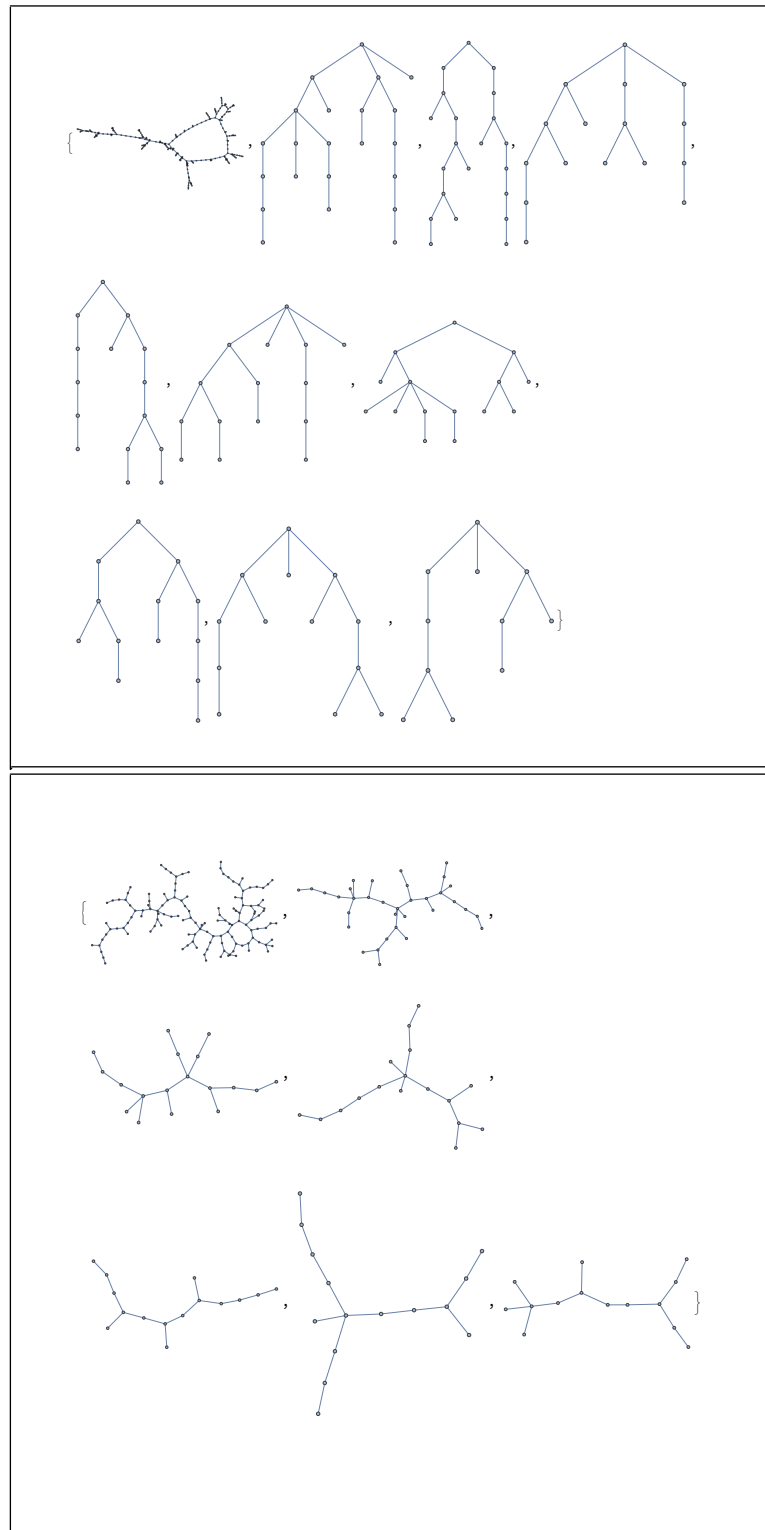


Figura 3: Dos redes de tipo Erdős-Rényi para $N = 1000$, $z = 0,98$ (arriba) y $z = 1,08$ (abajo). En el primer caso, el mayor agregado tiene 78 elementos, en el segundo 202. Para tener una imagen más limpia, se eliminaron en ambos casos los agregados de tamaño menor a 10.

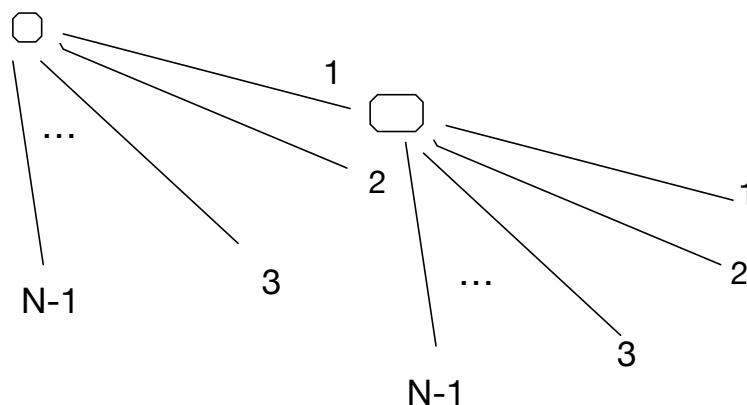


Figura 4: Ilustrando la estructura recursiva de la red de Erdős-Rényi.

tiende a cero. Por otro lado, si $z > 1$, el número de descendientes en cada generación crece exponencialmente y al final falla la aproximación que no hay ciclos. En este caso se conecta al infinito con probabilidad mayor que cero.

El método de recurrencia permite obtener muchos resultados de manera exacta. Por ejemplo, la probabilidad de conectarse al infinito, $P_\infty(z)$ se calcula de la manera siguiente: la probabilidad que un nodo dado se conecte al infinito se puede dividir entre una serie de casos exclusivos caracterizados por el número k de descendientes inmediatos del nodo: si el nodo tiene k descendientes, la probabilidad que a lo menos uno de estos descendientes se conecte al infinito es de

$$Q_{k,\infty}(z) = 1 - [1 - P_\infty(z)]^k \quad (6)$$

Por lo tanto, la probabilidad de conectarse al infinito es la suma, sobre todos los k , del producto de la probabilidad de tener k descendientes con la probabilidad $Q_{k,\infty}(z)$ que a lo menos uno de éstos se conecte al infinito. Pero la probabilidad de tener k descendientes está dada por (3), y obtenemos

$$P_\infty(z) = e^{-z} \sum_{k=1}^{\infty} \frac{z^k}{k!} Q_{k,\infty}(z). \quad (7)$$

Y esto, con algo de álgebra sencilla se reduce a:

$$P_\infty(z) = 1 - \exp[-zP_\infty(z)] \quad (8)$$

Esta ecuación no se puede resolver usando expresiones analíticas conocidas, pero sí se puede calcular con mucha facilidad. Asimismo se puede probar que, si $z \leq 1$, la única solución es cero, mientras en cuanto $z > 1$, existen 2 soluciones, una igual a cero y otra distinta. El modo de analizar la solución de (8) se muestra en la parte de arriba de la Figura 5, y el resultado del cálculo se ve en la parte baja. Vemos que la probabilidad de conectarse al infinito pasa *continuamente* de ser cero a ser positiva, y que cerca de la transición el crecimiento es lineal.

Los $n_s(z)$ también se evalúan de manera exacta. Contrariamente a lo que sucede en el caso de percolación en dimensión finita, para las redes de Erdős-Rényi existe una expresión sencilla válida para todas las s :

$$n_s(z) = \frac{s^{s-3}}{(s-1)!} z^{s-1} e^{-sz} \quad (9)$$

La prueba de ello sigue un camino análogo al del caso de $P_\infty(z)$, pero es algo más intrincada, y no es posible presentarla en este espacio.

Usando la fórmula de Stirling para $s!$ cuando $s \gg 1$ se obtiene que las cantidades $s^{5/2}n_s(z)$ caen sobre una sola curva si se reescala el eje s de manera apropiada. Esto lleva a la teoría del *escalamiento*, que es válida para la percolación en general, así como para otras transiciones de fase. Vemos en la Figura 6 el comportamiento de los $n_s(z)$ en función de s para determinados valores de z cercanos a, pero menores que, la transición en $z = 1$.

En resumen, vemos que las redes de Erdős-Rényi son un caso particular de la percolación, que se pueden describir de manera exacta y que corresponden *grosso modo* a la teoría del campo promedio de la percolación. Vale la pena destacar que es también posible estudiar las redes de Erdős-Rényi para el caso en que Np , en vez de ser una constante, diverge hacia infinito. En este caso, puede suceder que, con probabilidad 1, la red conste de una sola parte conexas. Estas redes son bastante más complicadas. En particular tienen una geometría mucho más compleja. No son árboles, y en particular contienen ciclos pequeños, como por ejemplo triángulos.

Esta ecuación no se puede resolver usando expresiones analíticas conocidas, pero sí se puede calcular con mucha facilidad. Asimismo se puede probar que, si $z \leq 1$, la única solución es cero, mientras en cuanto $z > 1$, existen 2 soluciones, una igual a cero y otra distinta. Esto se ve en la siguiente figura

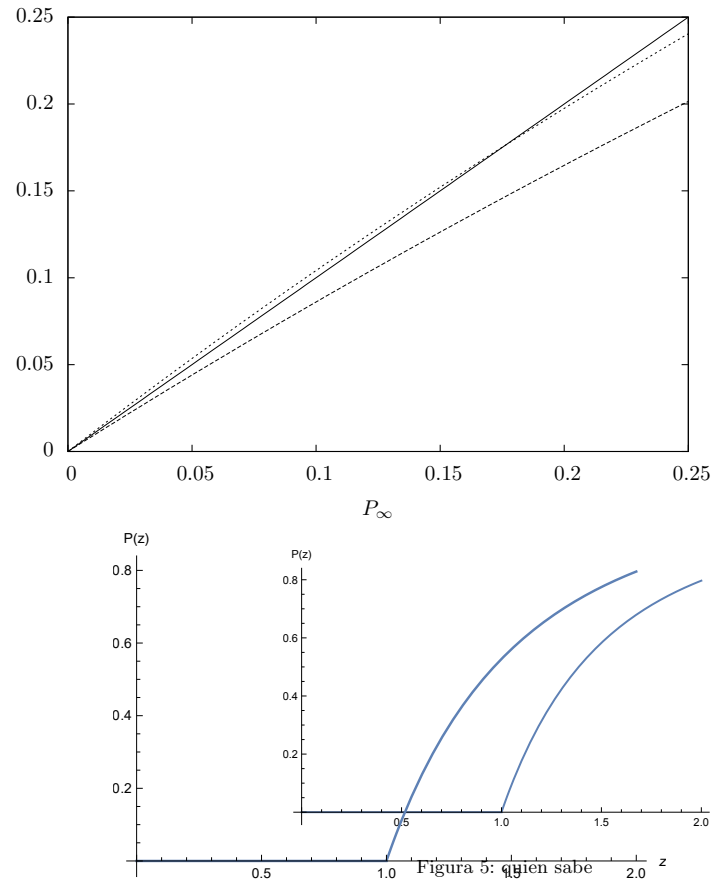


Figura 5: Arriba: esbozo de la solución gráfica de (8): las funciones $1 - \exp(-zP_\infty)$ se dibujan en función de P_∞ para dos valores de z : 0,9 y 1,1. Se dibuja también la función identidad (la recta de 45°). Vemos que la función cruza la recta en dos puntos sólo cruza si $z = 1$. Las curvas en un punto, sobre una, sólo si $z = 0,9$. Abajo: gráfico de la función $P(z)$. Esto lleva a la teoría del *escalamiento*, que es válida para la percolación en general, así como para otras transiciones de fase.

En resumen, vemos que las redes de Erdős-Rényi son un caso particular de la percolación, que se pueden describir de manera exacta y que corresponden *grosso modo* a la teoría del campo promedio de la percolación. Vale la pena destacar que es también posible estudiar las redes de Erdős-Rényi para el caso en que Np , en vez de ser una constante, diverge hacia infinito. En este caso, puede suceder que, con probabilidad 1, la red conste de una sola parte conexa. Estas redes son bastante más complicadas. En particular tienen una geometría mucho más compleja. No son árboles, y en particular contienen ciclos pequeños, como por ejemplo triángulos.

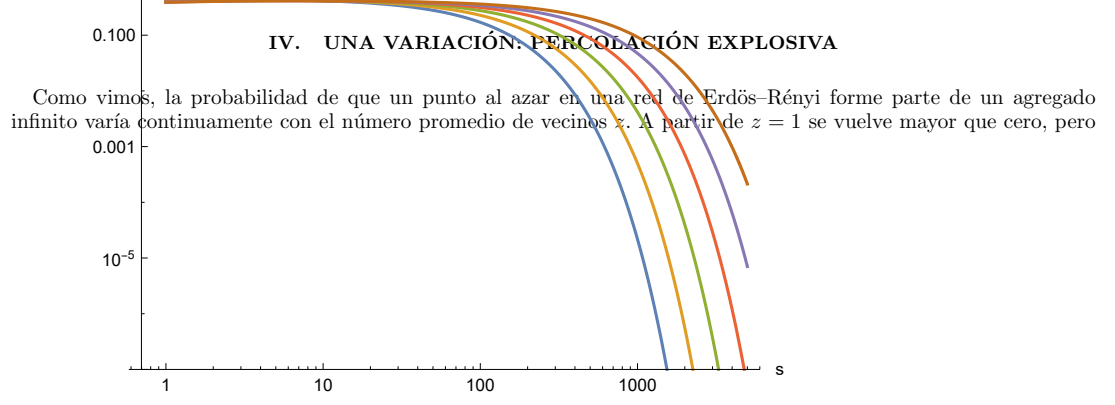


Figura 6: Se representan $s^{5/2}n_s(z)$ para varios valores de z cercanos a 1. Las curvas corresponden a los valores de z iguales a 0.865412, 0.887843, 0.906536, 0.922113, 0.935095, 0.945912.

IV. UNA VARIACIÓN: PERCOLACIÓN EXPLOSIVA

Como vimos, véase en particular la Figura 5, la probabilidad de que un punto elegido al azar en una red de Erdős–Rényi forme parte de un agregado infinito varía *continuamente* con el número promedio de vecinos z . A partir de $z = 1$ se vuelve mayor que cero, pero lo hace de manera continua.

En mecánica estadística existen dos tipos de transiciones de fase: algunas, como la percolación o las transiciones relacionadas con los puntos críticos, son continuas y pasan por una fase “autosimilar” en la que las correlaciones se extienden sobre longitudes muy grandes. A su lado existen—y son más comunes—las discontinuas, como la transición del agua al vapor: allí, la densidad cambia repentinamente en un factor de 1000. Por otro lado, en estos casos, no se pasa nunca por una fase de correlaciones de largo alcance, sino que ambas fases son “sencillas”, y una pasa de repente a ser más estable que la otra.

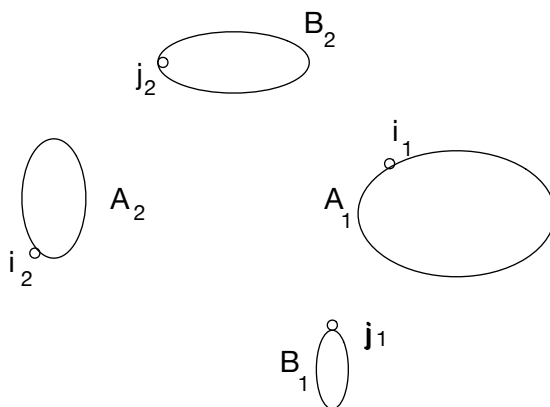


Figura 7: Descripción esquemática de la percolación explosiva: se escogen 2 pares de índices, $i_{1,2}$ y $j_{1,2}$

Siempre ha habido curiosidad acerca de la posibilidad de diseñar un modelo en redes o en percolación que mostrara tal cambio *discontinuo* en la probabilidad de pertenecer a una agregado infinito. Aquí presentaremos algo escuetamente un ejemplo que atrajo mucha atención por *casi* haberlo logrado. Este modelo se definió en [7], y se llama por ello a veces el modelo de Achlioptas.

Similarmente al modelo de crecimiento para las redes de Erdős–Rényi, empezamos por un conjunto de N puntos sin conexiones algunas. Describimos ahora como se desarrolla una conexión nueva, véase la Figura 7:

1. En cada paso, escoja 4 nodos al azar, en 2 parejas: digamos i_1, j_1 y i_2, j_2
2. Identifique los agregados A_1, B_1 que contienen con i_1, j_1 , y también para A_2, B_2
3. Identifique cuál de las dos parejas (A_1, B_1) y (A_2, B_2) lleva a *un crecimiento menor*, por ejemplo tomando el producto de los tamaños. Esta es la pareja que se conecta. La otra se descarta completamente.
4. Incrementar el tiempo en $1/N$.

En lo siguiente, para resaltar las similitudes con el modelo de Erdős–Rényi y la percolación, denotaremos el tiempo por p .

Como veremos, la numérica de este modelo parece dar con cierta claridad una transición discontinua. Veamos si podemos dar algún argumento que lo explique: el siguiente modelo es fácil de resolver, y se parece de alguna manera al anterior: a cada paso se buscan en todo el sistema el par de agregados más pequeños, y se juntan. Este modelo claramente tiene la característica del modelo de percolación explosiva que se busca limitar a cada paso el crecimiento, pero es mucho más exagerado.

Analicemos lo que sucede en este caso. Mientras haya agregados de tamaño 1, sólo se unen éstos, y los que se crean son de tamaño 2. Esto dura hasta el tiempo $1/2$. Después sólo se unen agregados de tamaño 2, y los que se crean son de tamaño 4. Esto dura por un tiempo de $1/4$, es decir, desde el tiempo $1/2$ hasta $3/4$. Siguiendo así vemos con facilidad que antes del tiempo 1 no aparece nunca un agregado infinito, pero al tiempo 1 aparece de manera repentina un agregado infinito que todo lo abarca.

¿Será este último modelo lo bastante similar al modelo de percolación explosiva, para que valga la misma conclusión? Veamos los resultados numéricos. Por un lado, para la distribución de tamaños de los agregados $n_s(t)$, observamos

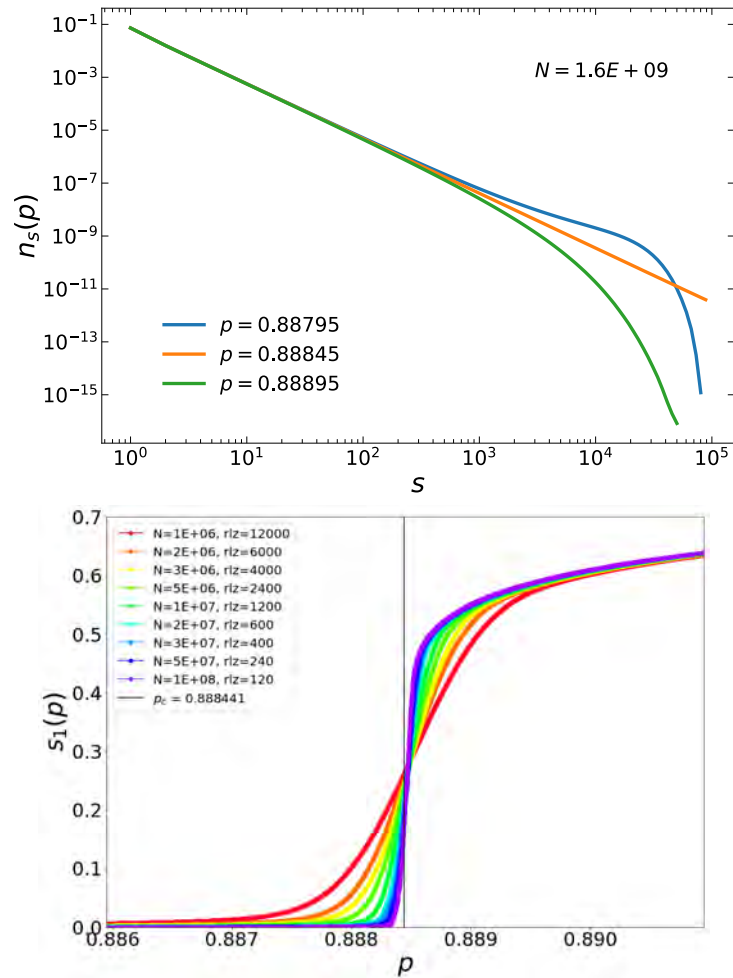


Figura 8: Arriba: gráfica doble logarítmica de $n_s(p)$ para valores de p tanto abajo, arriba e igual al valor crítico p_c , mostrando un claro comportamiento en ley de potencia. Abajo: tamaño del mayor agregado dividido por N , en función de p para varios valores del tamaño N del sistema, llegando hasta a 10^9 . El cambio se hace siempre más brusco, y no hay evidencia de una transición continua.

una ley de potencia cuando t toma el valor crítico, característica de las transiciones continuas, ya que sugiere una estructura auto-similar a todas escalas. Este argumento fue utilizado por da Costa *et al.* [13] para argumentar que la transición en efecto es continua. Mostramos esto en la parte de arriba de la Figura 8. Por otro lado, el tamaño del mayor agregado varía muy bruscamente, como también se evidencia en la Figuar 8, parte de abajo.

Para dilucidar esta duda, Grassberger [14] propuso ir más allá de mirar como varía el promedio del tamaño del mayor agregado, y propuso estudiar la *distribución de probabilidad* del mismo. En otras palabras se realizan muchas corridas y se paran al valor crítico de p . Se calcula en cada caso S_1 , el tamaño del agregado mayor dividido entre N y se dibuja el histograma correspondiente, es decir, para cada valor de S_1 se dibuja el número de corridas en las que se encontró. Esto se muestra en la Figura 9. Lo más notable de este resultado es el carácter *bimodal* de la distribución: lo que vemos es que al p crítico pueden pasar dos cosas: o se formó un agregado de tamaño proporcional a N , en cuyo caso el tamaño es una fracción notable de N (se trata de los máximos de las curvas de campana de la derecha, que se hallan alrededor de 0,4 para los mayores valores de N , y que son aún mayores si N es menor) o si no, no hay agregado de tamaño proporcional a N y el mayor agregado tiene un tamaño dividido por N muy cercano a cero, como lo muestran los picos a la izquierda.

Finalmente se reveló la solución del problema: en [15] Riordan y Warnke mostraron de manera rigurosa que la transición es continua. Específicamente, mostraron que si para un p el mayor agregado tiene un tamaño menor a \sqrt{N} , digamos, y si a otra $p' > p$ el agregado mayor tiene tamaño proporcional a N , entonces el número de pasos que separan estos dos estados debe ser proporcional a N . Por desgracia, el argumento que muestra esto no es lo bastante sencillo para poder presentarse aquí. Sin embargo, la rapidez de la transición como lo muestra la investigación numérica, no

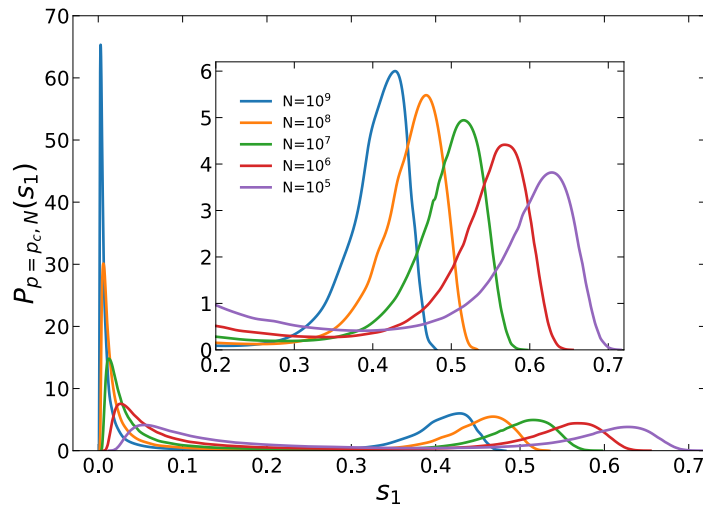


Figura 9: Histograma de tamaños del mayor agregado dividido por el número total de nodos N . El histograma se está tomando para el valor crítico de $p = 0,88845$. Vemos que hay dos posibilidades: a este valor de p puede haber un agregado del orden de N , o puede no haberlo. bajo estas circunstancias, el *promedio* no tiene mucho significado.

deja de ser asombrosa.

V. CONCLUSIONES

Hemos dado una visión muy de por encima de algunos sistemas en los que el concepto de conexión juega un papel importante. Conjuntos de puntos, llamados nodos, conectados por enlaces forman redes. Ejemplos de redes abundan: aeropuertos conectados por vuelos, usuarios de redes sociales unidos en “amistad”, nodos de la World Wide Web, todos estos sistemas y muchos más pueden verse desde el enfoque de estudiar la estructura de sus conexiones.

Vimos que se pueden describir varias propiedades usando sólo las conexiones de una red. Entre otras, podemos preguntar qué tan corto puede ser el camino que conecta dos puntos arbitrarios de la red. Esta propiedad es manifiestamente de importancia básica en la World Wide Web, pero también en otros contextos tiene relevancia. Otra propiedad importante es si el sistema está plenamente conectado o no. Y si no lo es, ¿cómo se distribuyen los tamaños de las componentes conexas? Finalmente, otra pregunta se refiere a la naturaleza de los nodos: ¿todos tendrán aproximadamente el mismo número de vecinos, o habrá nodos con un número cientos o miles de veces mayor que el promedio? Esta última posibilidad se realiza muchas veces en muchas redes en sistemas sociales: sea la red de aeropuertos, las redes sociales o las redes de citas a trabajos académicos.

Vimos dos modelos que evidencian ambas estas posibilidades: el modelo de crecimiento preferencial introducido por Barabási y Albert [5, 6] tienen la propiedad de tener algunos nodos con muy gran número de conexiones. Por otro lado, el modelo de Erdős–Rényi muestra una transición, como la que sucede en la percolación, entre un estado en que todas las partes conexas tienen un tamaño finito, a otro en el que existe una parte conexa cuyo tamaño crece con el tamaño del sistema. Ambos modelos tienen una definición muy sencilla. Esto, claro está, no nos debe llevar a creer que la compleja estructura de las redes del mundo real se deban a los mismos mecanismos. Pero sí nos indica que la existencia de complejidad no prueba necesariamente que haya sido el fruto de una compleja planeación: procesos aleatorios pueden en parte dar cuenta de ello.

Finalmente, vimos un modelo particular en el que lo que sucede aún no se entiende a cabalidad. La transición de un estado que sólo tienen agregados finitos, al otro en que existe un agregado macroscópico es, seguramente, continua, pero según el trabajo numérico que se ha realizado muestra una transición extremadamente brusca.

[1] Mandelbrot B.B. *The Fractal Geometry of Nature*, Time Books, second edition (1982)

[2] Barnsley M.F *Fractals Everywhere* second edition (2012) Dover

- [3] Coniglio, A., y Klein, W. (1980). “Clusters and Ising critical droplets: a renormalisation group approach”. J. Phys. A: Math. Gen., **13** (8), 2775
- [4] Milgram S. (1967) “The Small-World Problem” Psychology Today, **1** (1) 61–67
- [5] Barabási A.-L. y Albert R. (1999) “Emergence of scaling in random networks” Science **286** 509.
- [6] Albert R., Barabási A.-L. (2002). “Statistical mechanics of complex networks. Reviews of Modern Physics. **74** (1) 47–97.
- [7] Achlioptas D., D’Souza R.M., y Spencer J. (2009) Science **323**, 1453 .
- [8] Krapivsky, P.L., Redner S., y Leyvraz F. (2000) “Connectivity of growing random networks.” Phys. Rev. Lett. **85** (21) 4629.
- [9] Redner, S. (1998) “How popular is your paper? An empirical study of the citation distribution” The European Physical Journal B **4**, (2) 131–134.
- [10] Cohen R. y Havlin, S. (2003). “Scale-Free Networks Are Ultrasmall”. Phys. Rev. Lett. **90** (5) 058701.
- [11] Erdős, P.; Rényi, A. (1959). “On Random Graphs. I” Publicationes Mathematicae. **6** 290–297.
- [12] Stauffer D. y Aharony A. *Introduction to Percolation Theory*, second edition (2014) Taylor and Francis
- [13] da Costa R.A., Dorogovtsev S.N., Goltsev A.V., y Mendes J. F. F. (2010) “Explosive Percolation Transition is Actually Continuous” Phys. Rev. Lett. **105** 255701
- [14] Grassberger P., Christensen C., Bizhani G., Son S.-W., y Paczuski M. (2011) “Explosive Percolation is Continuous, but with Unusual Finite Size Behavior” Phys. Rev. Lett. **106** 225701
- [15] Riordan O. y Warnke L. (2011) “Explosive Percolation Is Continuous” Science **333** 322

Evolución temporal aproximada de un sistema optomecánico forzado con un sistema de dos niveles dentro de la cavidad

Notas basadas en la referencia: ENTROPY 2020, 22, 1373;
doi:10.3390/e22121373

L. Medina-Dozal, I Ramos-Prieto, J Récamier
Instituto de Ciencias Físicas, Universidad Nacional Autónoma de México,
Apdo. Postal 48-3, Cuernavaca, Morelos 62251, México.
E-mail: pepe@icf.unam.mx

Abstract

En estas notas presentamos un método aproximado para obtener el operador de evolución temporal correspondiente a un sistema optomecánico forzado con un sistema de dos niveles inmerso en la cavidad. El campo se acopla con el sistema de dos niveles mediante una interacción tipo Jaynes-Cummings y con el oscilador mecánico por presión de radiación. Los resultados obtenidos con nuestra aproximación son comparados con resultados puramente numéricos obtenidos resolviendo el problema con el Hamiltoniano completo. Para el conjunto de parámetros considerados en este trabajo encontramos un muy buen acuerdo entre ambos resultados.

1 Introducción

El Hamiltoniano de Jaynes-Cummings [1] modela uno de los sistemas cuánticos más sencillos, consiste en un modo cuantizado del campo electromagnético y un átomo de dos niveles. Debido a su importancia en óptica cuántica, este modelo ha sido estudiado por muchos autores tanto desde el punto de vista teórico como experimental [2, 3]. Es bien conocido que la evolución temporal del sistema depende drásticamente del estado inicial del campo de radiación, si el estado inicial es un estado coherente se presentan los fenómenos de colapsos y reavivamientos en la inversión atómica [4, 5] que son un resultado puramente cuántico. El modelo de Jaynes-Cummings ha sido generalizado en diversas formas, suponiendo que la interacción entre el campo y el átomo es no-lineal [6, 7, 8], generalizando el sistema atómico a un grupo de sistemas de dos niveles (modelo de Tavis-Cummings) [9], introduciendo un

medio no lineal en la propagación [10, 11] o incorporando simultáneamente un acoplamiento no lineal entre el átomo y el campo y un medio no lineal tipo Kerr [12].

Por otra parte, los sistemas optomecánicos proveen una herramienta para alcanzar y controlar el estado cuántico de objetos macroscópicos; lo hacen en rangos de frecuencia y masa que van desde Hz hasta GHz y desde 10^{-20} gramos hasta varios kilogramos. Los sistemas optomecánicos cuánticos permiten la detección de movimientos y fuerzas cercanas al límite fundamental impuesto por la mecánica cuántica [13, 14, 15]. El sistema optomecánico más sencillo, puede modelarse mediante una cavidad Fabry-Pérot con un espejo fijo y el otro extremo ligado armónicamente al cual se le permite oscilar bajo la acción de la presión de radiación del campo de luz de frecuencia ω_L presente en la cavidad. Gracias al forzamiento producido por la presión de radiación, la longitud de la cavidad se modifica y esto lleva a una modificación en la frecuencia de oscilación del espejo y a un amortiguamiento óptico en donde el campo óptico actúa como un fluido viscoso. Algunas aplicaciones recientes de este tipo de resonadores incluyen el enfriamiento de resonadores mecánicos para estudiar la transición entre el comportamiento clásico y cuántico [20] y la amplificación y medición de fuerzas a escalas nanométricas [21, 22].

La transferencia eficiente de información cuántica entre átomos y fotones es esencial y requiere de absorción y emisión controlada de fotones con una alta probabilidad. En [24] los autores lograron alcanzar el régimen de acoplamiento fuerte entre un oscilador mecánico y un átomo. El acoplamiento entre el movimiento de una membrana (que representa al oscilador mecánico) y el átomo está mediada por el campo cuantizado de luz en una cavidad de alta fineza.

En este trabajo obtenemos una aproximación para el operador de evolución temporal correspondiente a una cavidad optomecánica forzada así como el operador de evolución temporal para una interacción tipo Jaynes-Cummings, de esta manera encontramos un operador de evolución aproximado para el sistema híbrido.

2 Sistema optomecánico forzado

Consideremos un sistema descrito por un Hamiltoniano que consiste en un oscilador mecánico de frecuencia ω_m , un oscilador de campo de frecuencia ω_c un acoplamiento optomecánico debido a la presión de radiación y un

forzamiento del modo del campo con frecuencia ω_p :

$$\frac{\hat{H}_p}{\hbar} = \omega_c \hat{a}^\dagger \hat{a} + \omega_m \hat{b}^\dagger \hat{b} - G \hat{a}^\dagger \hat{a} (\hat{b} + \hat{b}^\dagger) + \Omega \cos(\omega_p t) (\hat{a} + \hat{a}^\dagger) \quad (1)$$

La intensidad del forzamiento Ω está relacionada con la potencia del láser P y la constante de acoplamiento optomecánico G está dada por

$$G = \frac{\omega_c}{L} \left(\frac{\hbar}{2m\omega_m} \right)^{1/2} \quad (2)$$

siendo L la longitud de la cavidad en ausencia de radiación y m la masa del oscilador mecánico. En ausencia del término de forzamiento (tomando $\Omega = 0$), el conjunto de operadores en el Hamiltoniano consta de $\{\hat{n}, \hat{N}, \hat{n}\hat{b}, \hat{n}\hat{b}^\dagger\}$ en donde hemos definido $\hat{n} = \hat{a}^\dagger \hat{a}$, $\hat{N} = \hat{b}^\dagger \hat{b}$. Usando que $[\hat{a}, \hat{a}^\dagger] = [\hat{b}, \hat{b}^\dagger] = 1$ y tomando los conmutadores entre los operadores del Hamiltoniano encontramos que si añadimos el operador \hat{n}^2 al conjunto, éste cierra ante la operación de conmutación. Esto es, podemos escribir el Hamiltoniano no forzado como:

$$\hat{H}_{np} = \sum_{i=1}^5 \phi_i \hat{X}_i \quad (3)$$

en donde $\hat{X}_1 = \hat{n}$, $\hat{X}_2 = \hat{N}$, $\hat{X}_3 = \hat{n}^2$, $\hat{X}_4 = \hat{n}\hat{b}$, $\hat{X}_5 = \hat{n}\hat{b}^\dagger$ y las ϕ_i se obtienen comparando las ecuaciones 1 y 3.

Debido a que hemos escrito el Hamiltoniano no forzado como una combinación lineal de operadores que cierran un álgebra de Lie, podemos utilizar el teorema de Wei-Norman [25, 26] y escribir el operador de evolución correspondiente de forma exacta como:

$$\hat{U}_{np}(t) = \prod_{i=1}^5 \exp(\alpha_i \hat{X}_i) \quad (4)$$

en donde las funciones α_i son funciones complejas, dependientes del tiempo, que satisfacen un conjunto de ecuaciones diferenciales acopladas, no lineales, ordinarias que se obtiene al sustituir la ecuación 4 en la ecuación de Schrödinger. Estas ecuaciones en general se resuelven de forma numérica salvo en casos particulares en los que se pueden encontrar soluciones analíticas. En este caso fue posible obtener las funciones $\alpha_i(t)$ de forma analítica [27].

El Hamiltoniano dado por la ecuación 1 puede escribirse como

$$\hat{H}_p = \hat{H}_{np} + \hbar \Omega \cos(\omega_p t) (\hat{a} + \hat{a}^\dagger) = \hat{H}_{np} + \hat{V}(t) \quad (5)$$

Escribimos el operador de evolución correspondiente a \hat{H}_p como $\hat{U}_p(t) = \hat{U}_{np}(t)\hat{U}_I(t)$ y sustituimos en la ecuación de Schrödinger

$$i\hbar \frac{\partial \hat{U}_p(t)}{\partial t} = \hat{H}_p \hat{U}_p(t), \quad \hat{U}_p(0) = \mathcal{I} \quad (6)$$

con $\hat{H}_p = \hat{H}_{np} + \hat{V}(t)$ obteniendo las ecuaciones

$$i\hbar \frac{\partial \hat{U}_{np}(t)}{\partial t} = \hat{H}_{np} \hat{U}_{np}(t), \quad \hat{U}_{np}(0) = \mathcal{I} \quad (7)$$

y

$$i\hbar \frac{\partial \hat{U}_I}{\partial t} = [\hat{U}_{np}^\dagger \hat{V}(t) \hat{U}_{np}] \hat{U}_I(t) = \hat{H}_I(t) \hat{U}_I(t), \quad \hat{U}_I(0) = \mathcal{I}. \quad (8)$$

El operador de evolución $\hat{U}_{np}(t)$ está dado por la ecuación 4 así que podemos construir de forma exacta el Hamiltoniano de interacción $\hat{H}_I(t)$ ya que el algebra de operadores es cerrada ante conmutación. Transformando los operadores \hat{a} y \hat{a}^\dagger obtenemos:

$$\begin{aligned} \hat{U}_{np}^\dagger \hat{a} \hat{U}_{np} &= e^{iE(t)(2\hat{n}+1)} e^{iF(t)[\hat{b}^\dagger e^{i\omega_m t/2} - \hat{b} e^{-i\omega_m t/2}]} \hat{a} e^{-i\omega_c t} \\ \hat{U}_{np}^\dagger \hat{a}^\dagger \hat{U}_{np} &= \hat{a}^\dagger e^{i\omega_c t} e^{-iF(t)[\hat{b}^\dagger e^{i\omega_m t/2} + \hat{b} e^{-i\omega_m t/2}]} e^{-iE(t)(2\hat{n}+1)} \end{aligned} \quad (9)$$

en donde las funciones $E(t)$ y $F(t)$ están dadas como

$$E(t) = 2 \left(\frac{G}{\omega_m} \right) \sin \left(\frac{\omega_m}{2} t \right) \quad (10)$$

$$F(t) = \left(\frac{G}{\omega_m} \right)^2 (\omega_m t - \sin(\omega_m t)) \quad (11)$$

ahora bien, el factor $G/\omega_m \ll 1$ [28] así que podemos aproximar

$$\hat{U}_{np}^\dagger \hat{a} \hat{U}_{np} \simeq \hat{a} e^{-i\omega_c t}, \quad \hat{U}_{np}^\dagger \hat{a}^\dagger \hat{U}_{np} \simeq \hat{a}^\dagger e^{i\omega_c t} \quad (12)$$

obteniendo $H_I(t) \simeq \hbar\Omega \cos(\omega_p t)(\hat{a} e^{-i\omega_c t} + \hat{a}^\dagger e^{i\omega_c t})$. El conjunto de operadores $\{\hat{a}, \hat{a}^\dagger, \mathcal{I}\}$ es cerrado ante conmutación y podemos usar el teorema de Weierstrass para escribir

$$U_I(t) = e^{\beta \hat{a}^\dagger} e^{\gamma \hat{a}} e^{\delta \mathcal{I}} \quad (13)$$

con las funciones β , γ , y δ tales que

$$\dot{\beta} = -i\Omega \cos(\omega_p t) e^{i\omega_c t} \quad (14)$$

$$\dot{\gamma} = -i\Omega \cos(\omega_p t) e^{-i\omega_c t} \quad (15)$$

$$\dot{\delta} = \beta \dot{\gamma} \quad (16)$$

con la condición inicial $\beta(0) = \gamma(0) = \delta(0) = 0$. Tenemos entonces una aproximación para el operador de evolución $U_p(t)$ expresada en términos de un producto de exponenciales por lo cual es relativamente fácil hacer la evolución temporal del estado inicial.

Considerando un estado inicial que es un producto de estados coherentes del campo y del oscilador mecánico $|\Psi(0)\rangle = |\alpha\rangle_c \otimes |\Gamma\rangle_m$ obtenemos el estado del sistema al tiempo t

$$\begin{aligned} |\Psi(t)\rangle &= \hat{U}_p(t)|\Psi(0)\rangle = \hat{U}_{np}(t)\hat{U}_I(t)|\Psi(0)\rangle \\ &= e^{\delta + \frac{1}{2}|\beta|^2 - \frac{1}{2}|\alpha + \beta|^2} e^{i\Im(\beta\alpha^*)} \sum_p \frac{(\alpha + \beta)^p}{\sqrt{p!}} e^{i(G/\omega_m)^2(\omega_m t - \sin \omega_m t)p^2} \\ &\times e^{-i[\omega_c t - \Im(\alpha_4 \Gamma^*)]p} |p\rangle_c \otimes |\Gamma_p(t)\rangle_m \end{aligned} \quad (17)$$

con $\Gamma_p(t) = \Gamma e^{-i\omega_m t} + p \frac{G}{\omega_m} (1 - e^{-i\omega_m t})$ de donde vemos que el enredamiento entre el campo y el oscilador mecánico es máximo para $\omega_m t = \pi$ y el sistema regresa a su estado no enredado para $\omega_m t = 2\pi$. Un estudio sobre la evolución temporal del número medio de fotones y fonones así como de la entropía lineal puede verse en la referencia [27].

3 Sistema optomecánico híbrido

El Hamiltoniano que describe el sistema optomecánico híbrido es:

$$\frac{\hat{H}}{\hbar} = \frac{\hat{H}_p}{\hbar} + \frac{\omega_a}{2} \hat{\sigma}_z + \lambda(\hat{a} \hat{\sigma}_+ + \hat{a}^\dagger \hat{\sigma}_-) \equiv \frac{\hat{H}_p}{\hbar} + \frac{\hat{W}}{\hbar} \quad (18)$$

en donde \hat{H}_p es el Hamiltoniano del sistema optomecánico forzado dado en la ecuación 1, ω_a es la frecuencia de transición atómica y λ es la constante de acoplamiento entre el campo y el sistema de dos niveles. Para obtener un operador de evolución aproximado para el sistema híbrido, escribimos

el operador de evolución como $\hat{U} = \hat{U}_p \hat{U}_{II}$ en donde \hat{U}_p es el operador de evolución obtenido en la sección anterior y $\hat{U}_{II}(t)$ satisface la ecuación:

$$i\hbar \frac{\partial \hat{U}_{II}(t)}{\partial t} = [\hat{U}_p^\dagger(t) W \hat{U}_p(t)] \hat{U}_{II}(t), \quad \hat{U}_{II}(0) = \mathcal{I}. \quad (19)$$

Al transformar la interacción obtenemos

$$\hat{H}_{II}(t) = \hat{U}_p^\dagger(t) W \hat{U}_p(t) \simeq \frac{\hbar\omega_a}{2} \hat{\sigma}_z + \hbar\lambda \left[(\hat{a} + \beta) \hat{\sigma}_+ e^{-i\omega_c t} + (\hat{a}^\dagger + \beta^*) e^{i\omega_c t} \right] \quad (20)$$

en donde de acuerdo con las aproximaciones usadas previamente, despreciamos términos proporcionales a G/ω_m y $(G/\omega_m)^2$ comparados con uno. Es conveniente escribir $\hat{H}_{II}(t) = \hat{H}_{II}^{(1)}(t) + \hat{H}_{II}^{(2)}(t)$ y $\hat{U}_{II}(t) = \hat{U}_{II}^{(1)}(t) \hat{U}_{II}^{(2)}(t)$ con

$$\hat{H}_{II}^{(1)}(t) = \frac{\hbar\omega_a}{2} \hat{\sigma}_z + \hbar\lambda (\beta e^{-i\omega_c t} \hat{\sigma}_+ + \beta^* e^{i\omega_c t} \hat{\sigma}_-) \quad (21)$$

$$\hat{H}_{II}^{(2)}(t) = \hbar\lambda (\hat{a} \hat{\sigma}_+ e^{-i\omega_c t} + \hat{a}^\dagger \hat{\sigma}_- e^{i\omega_c t}). \quad (22)$$

El Hamiltoniano $\hat{H}_{II}^{(1)}(t)$ es una combinación lineal de los operadores $\{\hat{\sigma}_+, \hat{\sigma}_-, \hat{\sigma}_z\}$ los cuales forman un conjunto cerrado ante conmutación y por tanto el operador de evolución temporal correspondiente está dado por:

$$\hat{U}_{II}^{(1)}(t) = e^{\alpha_z \hat{\sigma}_z} e^{\alpha_+ \hat{\sigma}_+} e^{\alpha_- \hat{\sigma}_-} \quad (23)$$

Al transformar $\hat{H}_{II}^{(2)}(t)$ con $\hat{U}_{II}^{(1)}(t)$ teniendo en cuenta que la constante de acoplamiento $\lambda \ll \omega_a$ y que $\omega_a \simeq \omega_c$ obtenemos:

$$[\hat{U}_{II}^{(1)\dagger}(t) \hat{H}_{II}^{(2)}(t) \hat{U}_{II}^{(1)}(t)] \simeq \hbar\lambda (\hat{a} \hat{\sigma}_+ e^{-i\omega_c t - 2\alpha_z} + \hat{a}^\dagger \hat{\sigma}_- e^{i\omega_c t + 2\alpha_z}). \quad (24)$$

Nótese que esta interacción es tipo Jaynes-Cummings y conserva el número total de excitaciones. Las funciones α_+ , α_- , y α_z satisfacen un conjunto de ecuaciones diferenciales ordinarias que se obtiene al sustituir $\hat{U}_{II}^{(1)}(t)$ en la ecuación de Schrödinger. Introducimos los operadores [8, 29, 30]

$$\hat{c} = \frac{1}{\sqrt{\hat{M}}} \hat{a} \hat{\sigma}_+, \quad \hat{c}^\dagger = \hat{a}^\dagger \hat{\sigma}_- \frac{1}{\sqrt{\hat{M}}} \quad (25)$$

con $\hat{M} = \hat{n} + \frac{1}{2}(1 + \hat{\sigma}_z)$ el número total de excitaciones en un escalón dado. Los estados que forman la base para el Hamiltoniano de Jaynes-Cummings son $\{|n, e\rangle, |n+1, g\rangle\}$ y corresponden al estado en que el campo tiene n fotones y

el átomo está en el estado excitado y al estado en el que el campo tiene $n + 1$ fotones y el átomo está en el estado base. El estado $|0, g\rangle$ en donde el campo no tiene fotones (está en el estado de vacío) y el átomo está en el estado base no se acopla con ningún otro estado. La acción de estos operadores sobre los estados de la base es:

$$\hat{c}|n, e\rangle = 0, \quad \hat{c}|n + 1, g\rangle = |n, e\rangle \quad (26)$$

$$\hat{c}^\dagger|n, e\rangle = |n + 1, g\rangle, \quad \hat{c}^\dagger|n + 1, g\rangle = 0 \quad (27)$$

$$\hat{M}|n, e\rangle = (n + 1)|n, e\rangle, \quad \hat{M}|n + 1, g\rangle = (n + 1)|n + 1, g\rangle \quad (28)$$

de las expresiones anteriores podemos obtener las relaciones de conmutación

$$[\hat{c}, \hat{c}^\dagger] = \hat{\sigma}_z, \quad [\hat{\sigma}_z, \hat{c}] = 2\hat{c}, \quad [\hat{\sigma}_z, \hat{c}^\dagger] = -2\hat{c}^\dagger \quad (29)$$

y $\hat{c}^2, \hat{c}^{\dagger 2}$ actuando sobre cualquier elemento de la base da cero. Podemos escribir el Hamiltoniano de interacción en términos de los operadores \hat{c}, \hat{c}^\dagger como

$$[\hat{U}_{II}^{(1)\dagger}(t)\hat{H}_{II}^{(2)}(t)\hat{U}_{II}^{(1)}(t)] = \hbar\lambda\sqrt{n+1}(\hat{c}e^{-i\omega_c t - 2\alpha_z} + \hat{c}^\dagger e^{i\omega_c t + \alpha_z})$$

y tenemos que resolver la ecuación

$$i\hbar\frac{\partial\hat{U}_{II}^{(2)}(t)}{\partial t} = \hbar\lambda\sqrt{n+1}(\hat{c}e^{-i\omega_c t - 2\alpha_z} + \hat{c}^\dagger e^{i\omega_c t + \alpha_z})\hat{U}_{II}^{(2)}(t) \quad (30)$$

cuya solución es

$$\hat{U}_{II}^{(2)}(t) = e^{\epsilon_1\hat{c}^\dagger}e^{\epsilon_2\hat{c}}e^{\epsilon_3\hat{\sigma}_z} \quad (31)$$

en donde una vez más utilizamos el teorema de Wei-Norman y las funciones ϵ_i se obtienen como en las ocasiones anteriores, resolviendo un conjunto de ecuaciones diferenciales ordinarias que se obtienen al sustituir $\hat{U}_{II}^{(2)}(t)$ en la ecuación de Schrödinger. Tenemos finalmente, una expresión aproximada para el operador de evolución temporal del sistema híbrido

$$\hat{U}(t) = \hat{U}_p(t)\hat{U}_{II}^{(1)}(t)\hat{U}_{II}^{(2)}(t) \quad (32)$$

en donde cada uno de los operadores ha sido escrito en términos de un producto de exponenciales y es relativamente fácil de aplicar al estado inicial del sistema.

4 Evaluación de observables

Consideremos un estado inicial dado por $|\Psi(0)\rangle = |n\rangle \otimes |e\rangle \otimes |\Gamma\rangle$ correspondiente a una cavidad con n fotones, un átomo en el estado excitado y un oscilador mecánico en un estado coherente Γ . Aplicando el operador $\hat{U}_{II}(t) = \hat{U}_{II}^{(1)}(t)\hat{U}_{II}^{(2)}(t)$ al estado inicial obtenemos:

$$\hat{U}_{II}^{(2)}(t)|n, e\rangle \otimes |\Gamma\rangle = e^{\epsilon_3} [|n, e\rangle + \epsilon_1|n+1, g\rangle] \otimes |\Gamma\rangle,$$

$$\begin{aligned} \hat{U}_{II}^{(1)}(t) [\hat{U}_{II}^{(2)}(t)|n, e\rangle \otimes |\Gamma\rangle] &= e^{\epsilon_3} [e^{\alpha z} (1 + \alpha_+ \alpha_-)|n, e\rangle + \alpha_- e^{-\alpha z}|n, g\rangle] \otimes |\Gamma\rangle \\ &+ e^{\epsilon_3} \epsilon_1 [e^{-\alpha z}|n+1, g\rangle + \alpha_+ e^{\alpha z}|n+1, e\rangle] \otimes |\Gamma\rangle, \end{aligned}$$

debido al término de forzamiento, el número total de excitaciones ya no es constante. Este estado puede ser escrito como

$$\hat{U}_{II}(t)|n, e\rangle \otimes |\Gamma\rangle = [c_1(t)|n, e\rangle + c_2(t)|n, g\rangle + c_3(t)|n+1, g\rangle + c_4(t)|n+1, e\rangle] \otimes |\Gamma\rangle \quad (33)$$

Si el estado del campo es descrito inicialmente por un estado coherente $|\alpha\rangle$ se obtiene

$$\begin{aligned} |\Psi(t)\rangle_I &= \hat{U}_{II}(t)|\alpha, e\rangle \otimes |\Gamma\rangle \\ &= \sum_{n=0}^{\infty} c_n [c_1(t)|n, e\rangle + c_2(t)|n, g\rangle + c_3(t)|n+1, g\rangle + c_4(t)|n+1, e\rangle] \otimes |\Gamma\rangle \end{aligned}$$

en donde $c_n = \exp[-\frac{1}{2}|\alpha|^2] \alpha^n / \sqrt{n!}$. El operador de evolución \hat{U}_p no involucra las variables atómicas, entonces podemos usar la ecuación anterior para calcular la probabilidad de encontrar al átomo en el estado excitado a un tiempo t arbitrario. El resultado es

$$P_e(\alpha, t) = |\langle e|\Psi(t)\rangle_I|^2 = \left| \sum_{n=0}^{\infty} c_n [c_1(t)|n\rangle + c_4(t)|n+1\rangle] \right|^2 \quad (34)$$

En la figura 1 se muestra la evolución temporal de la probabilidad de encontrar al átomo en el estado excitado cuando el estado inicial de la cavidad es un estado coherente con número medio de fotones $\langle \hat{n} \rangle = 4$, acoplamiento entre el átomo y la cavidad $\lambda/\omega_c = 0.0125$, la intensidad del forzamiento es $\Omega/\omega_c = 0.01$, la frecuencia de transición atómica es $\omega_a/\omega_c = 0.95$ y la

frecuencia del laser de forzamiento es $\omega_L/\omega_c = 0.5$. En ambos casos podemos observar el fenómeno de colapsos y reavivamientos, en el caso forzado la duración de los colapsos y la definición de los reavivamientos es mayor que en el caso no forzado y la probabilidad de encontrar al átomo en el estado excitado es también mayor.

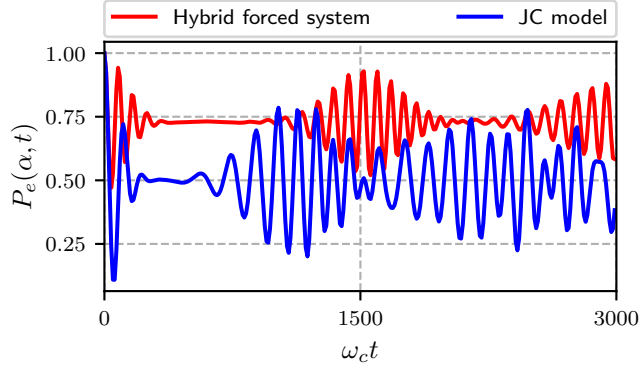


Figure 1: Probabilidad de encontrar al átomo en el estado excitado $P_e(\alpha, t)$ con $\alpha = 2$, $\omega_a/\omega_c = 0.95$, $\omega_L/\omega_c = 0.5$, $\lambda/\omega_c = 0.0125$, $\Omega/\omega_c = 0.01$. En rojo mostramos el caso para el sistema forzado, en azul el resultado obtenido con el modelo de Jaynes-Cummings

Calculemos ahora el valor medio del operador de número de fotones $\langle \hat{n}(t) \rangle$ y del operador de número de fonones $\langle \hat{N}(t) \rangle$. El número medio de fotones está dado por

$$\begin{aligned} \langle \hat{n}(t) \rangle &= \langle \Psi(0) | U_{II}^\dagger(t) \hat{U}_p^\dagger(t) \hat{n} \hat{U}_p \hat{U}_{II}(t) | \Psi(0) \rangle \\ &= {}_I \langle \Psi(t) | \hat{n}_I(t) | \Psi(t) \rangle_I \end{aligned} \quad (35)$$

en donde $\hat{n}_I(t)$ es el operador de número de fotones en la representación de interacción

$$\hat{n}_I(t) = \hat{n} + \beta^* \hat{a} + \beta \hat{a}^\dagger + |\beta|^2 \quad (36)$$

y para el número medio de fonones obtenemos

$$\hat{N}_I(t) = \hat{N} + (\alpha_4 \hat{b}^\dagger + \alpha_4^* \hat{b}) \hat{n}_I(t) + |\alpha_4|^2 \hat{n}_I^2(t), \quad (37)$$

Nótese que la evolución temporal del número de fonones depende de la evolución del número de fotones presentes en la cavidad. En la figura 2 se muestra la evolución temporal del número de fotones (panel izquierdo) y

fonones (panel derecho) para un estado inicial $|\Psi(0)\rangle = |\alpha\rangle \otimes |\Gamma\rangle \otimes |e\rangle$ en donde usamos los mismos parámetros que en la figura anterior y $\Gamma = 1$. Para verificar la validez de las aproximaciones hechas para llegar al operador de evolución dado en la Ec. 32 hicimos también un cálculo numérico resolviendo la ecuación de Schrödinger con el Hamiltoniano completo dado en la ecuación 18 utilizando el programa Python [31]. Es indiscutible la validez de las aproximaciones al menos en el rango de parámetros utilizados, parámetros que fueron obtenidos a partir de referencias de grupos experimentales [32]

Vemos en la figura que el número medio de fonones oscila con frecuencia ω_m partiendo de su valor inicial $N = 1$ hasta llegar aproximadamente a $N_{min} = 0.7$ de forma que el oscilador mecánico está enfriándose. El número medio de fotones muestra oscilaciones mucho más rápidas, de frecuencia ω_c debidas al intercambio entre el campo de la cavidad y el sistema atómico de dos niveles. Si hacemos la evolución a tiempos más largos, llegamos a observar la aparición de colapsos y reavivamientos similares a los observados en la figura 1.

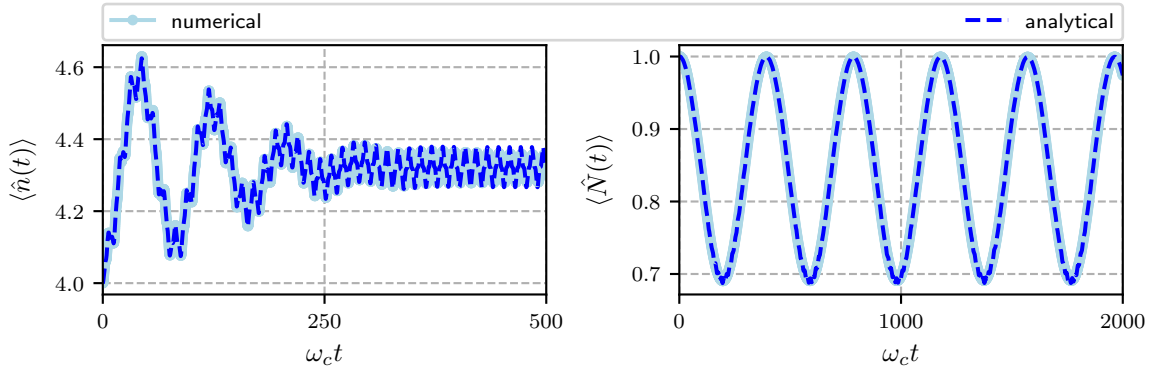


Figure 2: Evolución temporal del número medio de fotones (izquierda) y fonones (derecha). Resultados analíticos en azul fuerte, resultados numéricos en azul claro. Los parámetros del Hamiltoniano son $\alpha = 2$, $\omega_a/\omega_c = 0.95$, $\omega_L/\omega_c = 0.5$, $\omega_m/\omega_c = 0.016$, $G/\omega_c = 0.00032$, $\lambda/\omega_c = 0.0125$ y $\Omega/\omega_c = 0.01$.

5 Conclusiones

En este trabajo presentamos un método aproximado para construir el operador de evolución temporal de un sistema híbrido consistente en un sistema optomecánico forzado y un sistema atómico de dos niveles inmerso en la cavidad. El átomo se acopla solamente con la cavidad mediante una interacción tipo Jaynes-Cummings. Para resolver el problema separamos el Hamiltoniano en dos partes, una corresponde a un sistema optomecánico forzado y la otra a un átomo libre y una interacción tipo JC. El operador de evolución para el sistema optomecánico forzado es escrito en forma de un producto de exponenciales y se utiliza para generar un Hamiltoniano de interacción en una representación de interacción generalizada. Linearizando este Hamiltoniano se obtiene el correspondiente operador de evolución. Conocido el operador de evolución calculamos la evolución temporal de varias observables físicas como son la probabilidad de encontrar el átomo en el estado excitado y la evolución temporal de los valores medios de fotones y fonones. Comparamos nuestros resultados con resultados numéricos precisos y encontramos una excelente concordancia entre ellos.

Agradecimientos: Agradecemos a Reyes García por el mantenimiento a nuestras computadoras. Agradecemos apoyo parcial de la Dirección General de Asuntos del Personal Académico (DGAPA-UNAM) mediante el proyecto IN1111119. I. Ramos-Prieto agradece apoyo posdoctoral otorgado por DGAPA-UNAM.

References

- [1] Jaynes E. T. and Cummings F. W., 1963 Proc. IEEE **51** 89.
- [2] Shore B and Knight P L 1993, J Mod. Opt. **40**, 1195-238
- [3] Andrew D Greentree, Jens Koch and Jonas Larson, Journal of Physics B: Atomic, Molecular and Optical Physics **46**, 220201 (2013).
- [4] Rempe G. and Walther H., 1987, Phys. Rev. Lett. **58** 353.
- [5] Haroche S., Raymond J., Exploring the Quantum: Atoms, cavities and photons, Oxford University Press 2006

- [6] Buck B and Sukumar C. V., 1981 Phys. Lett. A **81** 132
- [7] Buzek V. 1989 Phys. Rev. A **39** 3196
- [8] Cordero S and Récamier J., 2011 J. Phys. B: At. Mol. Opt. Phys. **44** 135502
- [9] Tavis M and Cummings F W., 1968 Phys. Rev. **170** 379
- [10] Agarwal G. S. and Puri R. R., 1989 Phys. Rev. A **39** 2969
- [11] Werner M. J. and Risken H., 1991 Quantum Opt. **3** 185
- [12] O de los Santos-Sánchez and J Récamier, 2012, J. Phys. B: At. Mol. Opt. Phys. **45**, 015502
- [13] Aspelmeyer M., Meystre P. and Schwab K., 2012 Quantum optomechanics, Phys. Today **65** 29-35
- [14] Kippenberg T. J. and Vahala K., 2008 Cavity optomechanics: back-action at the mesoscale, Science **321** 1172-1176
- [15] P Meystre, Ann. Phys. (Berlin) **525**, No. 3, 215-233 (2013)
- [16] Corbitt T., Chen Y., Innerhofer E., Müller-Ebhardt H, et. al., 2007 Phys. Rev. Lett. **98** 150802
- [17] Metzger C. H., and Karrai K., 2004 Nature **432** 1002
- [18] Thompson J. D., Zwickl B. M., Jayich A. M. Marquardt F., Girvin S. M., and Harris J.G., 2008 Nature **452** 72
- [19] Corbitt T., and Mavalvala N., 2004 J. Opt. B **6** S675-83
- [20] Gigan S., Böhm H. R., Paternostro M., Blaser F., Langer G., et. al., 2006 Nature **444** 67
- [21] Carmon T., Rokhsari H., Yang L., Kippenberg T. J., and Vahala K. J., 2005 Phys. Rev. Lett. **94** 223902
- [22] Metzger C., Ludwig M., Neuenhahn C., et. al., 2008 Phys. Rev. Lett. **101** 133903

- [23] G. T. Moore, "Quantum theory of the electromagnetic field in a variable length one-dimensional cavity", *J. Math. Phys.* **11** 2679-2691 (1970).
- [24] Hammerer, K.; Wallquist, M.; Genes, C.; Ludwig, M.; Marquardt, F.; Treutlein, P.; Zoller, P.; Ye, J.; Kimble, H. J.; , "Strong coupling of a mechanical oscillator and a single atom", *Phys. Rev. Lett.* **103**, 063005 (2009).
- [25] Wei J and Norman E., "On global representations of the solutions of linear differential equations as a product of exponentials", *Proc. Am. Math. Soc.* **15**, 327-334 (1964).
- [26] Mancini S., Man'ko V I, Tombesi P, "Ponderomotive control of quantum microscopic coherence", *Phys. Rev. A* **55** 3042-3050 (1997).
- [27] A Paredes-Juárez, I Ramos-Prieto, M Berrondo and J Récamier, "Lie algebraic approach to quantum driven optomechanics", *Phys. Scr.* **95** 035103 (2020).
- [28] Ventura-Velázquez C., Rodríguez-Lara B., and Moya-Cessa H., "Operator approach to quantum optomechanics", *Phys. Scr.* **90** 068010 (2015).
- [29] Rodríguez-Lara, BM.; Moya-Cessa, H.; "The exact solution of generalized Dicke models via Susskind-Glogoweer operators", *J. Phys. A: Math. Theor.* **46** 095301 (2013).
- [30] Ramos-Prieto I.; Paredes A.; Récamier J.; Moya-Cessa H.; "Approximate evolution for a system composed by two coupled Jaynes-Cummings Hamiltonians", *Phys. Scr.*, **95**, 034008 (2020).
- [31] Johansson, J.R.; Nation, P.D.; Nori, F., "QuTip: An open-source Python framework for the dynamics of open quantum systems", *Comput. Phys. Commun.* **2012**, 183, 1760-1772.
- [32] Aspelmeyer, M.; Kippenberg, T.J.; Marquardt, F., "Cavity optomechanics", *Rev. Mod. Phys.* 2014, **86**, 1391-1452.

Superfluidez en redes ópticas

R. Paredes¹

¹*Instituto de Física, Universidad Nacional Autónoma de México,
Apartado Postal 20-364, México D.F. 01000, México.**

En el presente capítulo se presenta un resumen del fenómeno de superfluidez en gases atómicos ultrafríos confinados en redes ópticas. Primero se hace referencia al descubrimiento de la superfluidez en átomos de helio a temperatura de 2.17 K y sus características principales. Posteriormente se describe la ocurrencia del mismo fenómeno, pero en átomos alcalinos a temperaturas de nK. Antes de hablar de la superfluidez en estos átomos, brevemente se discute el concepto de temperatura y la forma en la que entre 1980 y 1995 se consiguieron temperaturas alrededor de 150 nK en gases atómicos ultrafríos. En los años subsecuentes se demostró que en dichas circunstancias el gas de átomos alcalinos se comporta como un superfluido, siendo la existencia de vorticidad un elemento contundente. Finalmente se describe cómo en redes ópticas, es decir en arreglos de luz con simetría de traslación, también se manifiesta el fenómeno de superfluidez en gases atómicos ultrafríos.

I. INTRODUCTION

En 1937 Donald Misener, John Allen y Piotr Kapitsa descubrieron que a una temperatura de 2.17 K el ^4He en su estado líquido exhibe un comportamiento muy peculiar caracterizado por la ausencia súbita y total de viscosidad. El estudio particular que condujo a este hallazgo fue el abrupto cambio que tiene la viscosidad cuando ^4He en su estado líquido fluye a través de tubos muy angostos de radio entre $10^{-5} - 10^{-2}$ cm a una temperatura de 2.17 K, se vuelve cero repentinamente!. No solamente el comportamiento de la viscosidad evidenció la existencia de un nuevo estado de la materia, también la capacidad de fluir sin resistencia en el caso del helio líquido se manifestó en el famoso efecto de fuente y al escapar de un contenedor subiendo por las paredes del mismo, venciendo la gravedad. De forma más técnica la capacidad calorífica demostró también la transición.

Sin duda el logro de Heike Kamerlingh Onnes en 1908 al conseguir la licuefacción de helio a una temperatura de unos cuantos K, fue el paso decisivo para el descubrimiento del fenómeno de superfluidez. Sin embargo, la carrera por el enfriamiento, es decir la capacidad de reducir la temperatura en un material, no se detuvo a inicios del siglo XX, sino que continuó por las siguientes décadas. Sin bien el establecimiento de la termodinámica y particularmente la el conocimiento experimental y fenomenológico de sustancias puras permitieron conseguir temperaturas cercanas al cero absoluto, fue necesario otro gran paso para reducir la temperatura aún más y acercarse a los -273 K. Precisamente a inicios del siglo XX una nueva disciplina de la ciencia, la mecánica cuántica, al tiempo que cambiar la concepción del comportamiento de la materia, abrió el camino que condujo a alcanzar de temperaturas cercanas a una centena de nK. Este gran logro, dió lugar en 1995 a la observación experimental de una fase de la materia predicha en 1925 por Satyendra Nath Bose y Albert Einstein, la llamada condensación de Bose-Einstein, que se caracteriza por la ocupación macroscópica del estado base, es decir una fase de la materia en la que un número gigantesco de átomos ocupan el estado de menor energía. Más tarde, en 1999 se consiguió, también en su estado degenerado, confinar un gas de Fermi. Ambos acontecimientos constituyen la primera realización experimental de la materia en su estado degenerado. Vale la pena resaltar en este punto que si bien la condensación de Bose-Einstein y el gas degenerado de Fermi se refieren a gases ideales, es decir materia constituida por átomos que no interactúan entre sí, las observaciones en el laboratorio en 1995 y 1999 corresponden a gases altamente diluidos, en los que solamente las interacciones entre pares de átomos son relevantes. Fue determinante la consecución de la condensación de Bose-Einstein y el gas degenerado de Fermi para el ulterior uso de los mismos para crear fases de la materia que solamente se presentan en el régimen de degeneración cuántica.

En estas notas se describe la ruta desde el suceso de la licuefacción de gases nobles, pasando por el elemento crucial que representó el conocimiento de la materia a través de las leyes de la mecánica cuántica, y que permitió usar la física atómica para reducir la temperatura de un gas atómico, hasta la manipulación y control de grandes conglomerados de átomos que colectivamente, y en su estado degenerado, son los ingredientes esenciales de donde emergen diversos estados de la materia. En particular el interés en las siguientes secciones es describir cómo se observó la superfluidez en redes ópticas en gases moleculares ultrafríos y mencionar en forma sucinta uno de los esquemas que permiten describir dicho fenómeno en forma teórica.

*Electronic address: rosario@fisica.unam.mx

II. TEMPERATURA

En esta sección se describe primeramente el contexto que en el que ocurrió el descubrimiento de la superfluidez en ^4He , y posteriormente el gran paso que fue lograr el estado superfluido, pero en gases de átomos alcalinos, a una temperatura mucho menor. La fase superfluida en este caso ocurrió en la llamada condensación de Bose-Einstein, que se consiguió por primera en 1995 en vapores de átomos de sodio y rubidio. Para entender cómo se logró observar estas fases, primero es necesario hacer referencia al concepto de temperatura.

Técnicamente la temperatura es proporcional a la energía cinética de los átomos y moléculas que componen un cuerpo, o dicho en otras palabras, la temperatura es una medida de la velocidad de los constituyentes de la materia. Sin duda, la temperatura guarda una estrecha relación con el calor suministrado a un cuerpo. Cuanto más calor tiene un cuerpo, mayor será su temperatura, la pregunta entonces es, ¿cómo se puede “quitar” calor a un cuerpo?, específicamente al conjunto macroscópico de átomos que componen a un gas. En ese sentido se dice que bajar la temperatura de un cuerpo significa conseguir que el inmenso conjunto de átomos que lo componen reduzcan su velocidad. La tercera ley de la termodinámica establece que el cero absoluto es inalcanzable, es decir un estado de la materia en el que los átomos y moléculas de una sustancia se hallen en reposo absoluto, sin embargo nada prohíbe acercarse arbitrariamente al cero absoluto. Las menores temperaturas alcanzadas hacia principios del siglo XX, con refinados conocimientos empíricos en el marco de la termodinámica, se consiguieron en distintos elementos. Por ejemplo se logró la licuefacción de oxígeno, nitrógeno, hidrógeno y helio a temperaturas de 90 K, 77 K, 20 K y 4.2 K respectivamente. Gracias al hecho de contar con helio a esas temperaturas fue que se descubrió la superconductividad en 1911. Más adelante, al hablar de la superfluidez del helio, se hará referencia a las similitudes que comparten los fenómenos de superfluidez y superconductividad.

Volviendo a la pregunta de cómo reducir la temperatura de un cuerpo, es inevitablemente necesario decir que la mecánica cuántica, y particularmente el entendimiento de que las leyes que rigen el comportamiento de la materia, que nacieron a finales del siglo XIX, fueron cruciales para “establecer un canal de comunicación” con los átomos y lograr que un conjunto macroscópico de ellos reduzca su velocidad. La clave de esa *comunicación* es que el intercambio de energía con los átomos ocurre a través de la luz. Como establece la teoría de Bohr, que describe al átomo de hidrógeno a través de un modelo semiclásico, los electrones van de un nivel de energía a otro vía la absorción o emisión de fotones, que, como sabemos son los cuantos de luz. Estos fotones son precisamente el canal de comunicación con la materia. Específicamente, las técnicas de enfriamiento láser y enfriamiento evaporativo se basan en utilizar en el forma apropiada la absorción de luz láser, sintonizada con las frecuencias naturales de las transiciones electrónicas, para conseguir que después de repetidos ciclos, los átomos de una nube atómica evolucionen, vía colisiones, a un gas cuya temperatura gradualmente disminuye, este proceso se llama la “rampa de enfriamiento”. Si bien no es la intención de las presentes notas el proveer de una descripción detallada de cómo es que operan las citadas técnicas, en los siguientes dos párrafos se describe en forma breve la esencia de las mismas.

La forma más sencilla de explicar la técnica de enfriamiento láser es considerando un solo átomo de masa m viajando de derecha a izquierda a una cierta velocidad v y momento $p = mv$. Si se hace incidir un haz de luz láser de izquierda a derecha, tal que la frecuencia de los fotones del láser coincide con la frecuencia de la transición electrónica del átomo, éste absorberá uno de estos fotones. El momento del átomo $p = mv$ se reducirá asociado al hecho que el momento del fotón es $-h/\lambda$, así, posterior a la absorción del fotón se tendrá que el átomo cambió su momento en una cantidad $p - h/\lambda$. Un tiempo después el fotón absorbido será reemitido dejando al átomo nuevamente como lo encontró antes del que fuera absorbido. Parecería que dicho proceso no es eficaz en la reducción de la velocidad, sin embargo, lo es pues la emisión del fotón absorbido ocurre en una dirección arbitraria y no precisamente en la dirección contraria en la que. Después de un miles de veces que ocurre el proceso de absorción unidireccional y emisión en dirección arbitraria, el átomo reduce su velocidad. En números explícitos se tiene que a una temperatura de 300 K un átomo viaja a una velocidad promedio de 570 m/s. Para un átomo de sodio cuya línea de emisión típica es $\lambda = 589$ nm, el cambio de momento después de que un fotón ha sido absorbido es de 1.12×10^{-27} kg m/s, lo que da lugar a una reducción de la velocidad del átomo de 3 cm/s. Se requieren aproximadamente 2×10^4 absorciones para frenar un átomo. En la figura 1 se ilustra esquemáticamente la técnica de enfriamiento láser.

Ciertamente es necesario mencionar que el efecto Doppler juega en papel importante en el proceso de enfriamiento láser, y también que hay un límite inferior de la temperatura que se puede alcanzar a través de esta técnica. No es posible reducir la temperatura de un gas de átomos alcalinos por debajo de los μK . Es en ese momento cuando entra en juego la técnica de enfriamiento por evaporación. En palabras simples, esta técnica es un proceso selectivo que consiste en “retirar” de la nube del gas a los átomos más calientes, de manera que a través de colisiones la nube remanente reduzca su temperatura. Es importante decir que la forma en la que los átomos más calientes son retirados de la nube, es haciendo incidir un pulso de radio frecuencia, también mencionar que a medida que el gas se va enfriando en etapas sucesivas de estados de equilibrio, hay una reducción inevitable de átomos en la nube térmica. Lo importante es que aún esa nube posee una densidad que permite alcanzar el estado de condensación. Las dos técnicas, de enfriamiento láser y enfriamiento evaporativo, fueron claves en la observación de la fase condensada en un gas de 5×10^5 átomos de

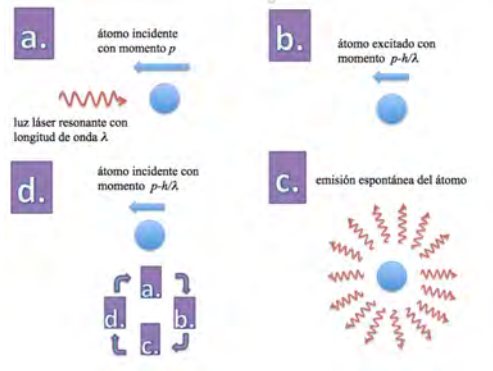


FIG. 1: Figura esquemática del proceso de pérdida de velocidad/momento de los átomos a través de la técnica de enfriamiento láser.

Na a densidades que superan 10^{14} cm^{-3} a una temperatura de $2 \mu\text{K}$ [1].

III. SUPERFLUIDEZ

La superfluidez es un estado de la materia caracterizado por la ausencia total de viscosidad. Como se ha mencionado antes, fue observada primeramente en ^4He líquido a una temperatura de 2.17 K. Vale la pena resaltar que este elemento químico está constituido de dos protones, dos neutrones y dos electrones, y por ello sabemos que es un átomo de tipo bosónico. En la figura 2 se observa el diagrama de fases del ^4He , en el mismo se pueden distinguir las fronteras, en el espacio presión-temperatura, que dividen las fases sólida, líquida, gaseosa y superfluida. La fase superfluida es también conocida como He-II, y como se aprecia de la figura 2 se restringe a una zona bien definida del espacio presión-temperatura. Años más tarde, también fue posible observar la fase superfluida en un isótopo de helio, el ^3He , en este caso los átomos poseen dos protones, un neutrón y dos electrones, ciertamente este isótopo del helio no es abundante y, de hecho, debe ser producido en el laboratorio. Pevio al descubrimiento de la fase superfluida en ^4He ,

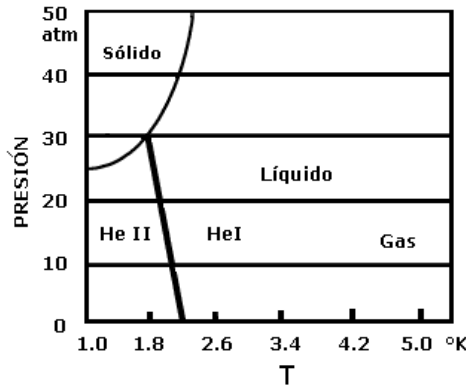


FIG. 2: Diagrama de fases de ^4He . Comúnmente la fase superfluida del ^4He es identificada como He-II, en tanto que la fase no superfluida, o normal es conocida como He-I.

y derivado de la posibilidad de contar con un baño térmico para poder enfriar, es decir un líquido a una temperatura cercana al cero absoluto, fue posible estudiar las propiedades de la materia en el régimen de bajas temperaturas. Probablemente el hecho de mayor trascendencia en este respecto fue el descubrimiento de la superconductividad en metales. En 1911 Heike Kamerlingh Onnes se encontraba estudiando la conductividad del mercurio cuando descubrió que a 4.2 K la resistencia eléctrica se anula por completo. Aún en ausencia de un campo, la corriente eléctrica persiste en el mercurio por debajo de los 4.2 K. En la Fig. 3 se aprecia la caída abrupta a una temperatura particular, identificada como temperatura crítica T_c , tanto de la resistividad ρ , como del calor específico C_v . Dicha temperatura es precisamente 4.2 K. Lo esperado dentro de la teoría del electromagnetismo es que la resistencia de los metales

disminuya linealmente con la temperatura a medida que esta tiende a cero. El que no ocurra de esta forma hace ver que los átomos que componen la materia manifiestan un comportamiento peculiar cuando se atraviesa cierto umbral de temperatura.

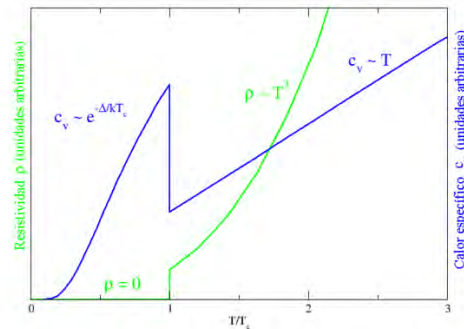


FIG. 3: Resistividad y calor específico del mercurio como función de la temperatura. La temperatura crítica identificada como T_c en la figura, tiene un valor de 4.2 K.

En este punto vale la pena enfatizar que las fases superfluida y superconductora se comportan en forma análoga en el siguiente sentido. Mientras que los electrones en el mercurio fluyen sin resistencia a través de la red, los átomos neutros de helio fluyen también sin ninguna resistencia cuando el líquido es puesto en un contenedor. Además, “curiosamente” ambas fases se presentan a muy bajas temperaturas cuando los átomos que de mercurio y helio en cada caso, siguen las leyes de la mecánica cuántica. Podríamos decir que el hecho contundente es que en el régimen de degeneración cuántica, y como resultado de que un número macroscópico de átomos es sometido a dicha condición, emergen nuevas fases de la materia, la superfluidez y la superconductividad. Otra forma de decirlo es señalando que las transiciones a los estados superfluido y superconductor son manifestaciones del comportamiento colectivo de la materia en el régimen cuántico.

La existencia de la fase supefluida en ^4He fue explicada en 1947 por Nikolai Bogoliubov considerando como elemento base, la interacción entre pares de bosones en un sistema constituido por un número macroscópico de ellos, y su estado de menor energía. Por otro lado, el fenómeno de superconductividad descubierto en 1911 fue explicado por Bardeen, Cooper y Schrieffer en la llamada “Teoría de la Superconductividad” en un artículo publicado en 1957 [2]. Su teoría, conocida como BCS, se basa en dos aspectos cruciales, uno que involucra la física de los portadores de carga y otro relacionado con la forma en la que se comporta colectivamente la materia constituida por esos portadores.

Dada la analogía entre los estados superfluido y superconductor, se especula que el entendimiento de los átomos ultrafríos, que también son superfluidos, abre un panorama para acceder al entendimiento de la llamada superconductividad de alta temperatura, descubierta en 1986 por Bednorz y Müller en materiales cerámicos (también denominados cupratos por su composición a base de cobre, YBaCuO) a una temperatura superior a los 90 K. A la fecha se conoce gran diversidad de compuestos que exhiben este fenómeno; de hecho, la temperatura crítica más alta reportada para aleaciones de bario y cobre se encuentra entre 133 -138 K. Hasta el momento no ha sido explicada, la llamada superconductividad de alta temperatura. Así, el simular la materia en el régimen cuántico, ya sea a través modelos teóricos o mediante su implementación controlada en un laboratorio, son dos valiosas herramientas que van de la mano en el objetivo de entender las transiciones de fase a los estados superfluido y superconductor de alta temperatura. Del lado teórico se pueden predecir fenómenos susceptibles de ser medidos en el laboratorio, y viceversa, la contraparte experimental permite delinear el modelo teórico pertinente para describir la física dentro de un laboratorio.

IV. SUPERFLUIDEZ EN REDES ÓPTICAS

Posterior al logro de la condensación de Bose-Einstein en vapores de átomos alcalinos en 1995, surgieron nuevos campos de estudio tanto teóricos como experimentales, particularmente en estas notas se hace referencia a la exploración de fases de la materia como resultado de la capacidad de emular el comportamiento de electrones que se mueven en un potencial cristalino. Antes de describir cómo es que este tipo de arreglos son implementados en un laboratorio es pertinente aclarar que los condensados de Bose-Einstein, y en general todos los gases atómicos ultrafríos son producidos en campos electromagnéticos, que en buena forma están bien descritos por un potencial de tipo armónico. En general, una forma de controlar la dimensión efectiva en la que los átomos se mueven es variando el orden de magnitud de la frecuencia de la trampa armónica originalmente isotrópica. Ello da lugar a que en la dirección o direcciones

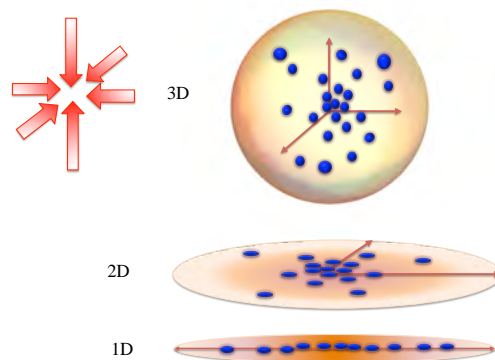


FIG. 4: Figura ilustrativa de átomos ultrafríos moviéndose en 3, 2 y 1 dimensión.

en las que la frecuencia aumenta el movimiento se ve drásticamente reducido. Así, en lugar de que los átomos se muevan en las tres direcciones espaciales, su movimiento ocurrirá en la dirección o direcciones en la que la frecuencia es menor. En la figura 4 se ilustra el espacio efectivo en el que se mueven los átomos del gas ultrafrío.

La posibilidad de contar con un “laboratorio” de átomos neutros cuyas interacciones entre ellos pueden ser controladas externamente, más el hecho que estos conglomerados macroscópicos constituyentes del gas neutro puedan moverse a través de arreglos de luz con simetría de traslación, hace que dichos sistemas lleven el nombre de simuladores cuánticos. En otras palabras, hay dos ingredientes fundamentales que dan como resultado que los átomos neutros del gas ultrafrío se comporten de forma análoga a los electrones de un sólido: la capacidad de “sintonizar” las interacciones entre los átomos, y el potencial de confinamiento que forma una red óptica. Es en sí un tema completo la descripción de la forma en la que vía campos magnéticos externos, se pueden controlar las interacciones entre átomos neutros. Estas notas se restringen a enunciar la forma en la que los gases ultrafríos se confinan en redes ópticas o cristales de luz.

Las llamadas las redes ópticas son resultado de la superposición de haces de luz láser formando una onda estacionaria. El ejemplo más sencillo ocurre en una dimensión, dos haces de luz láser propagándose en sentidos contrarios forman una onda estacionaria cuyos mínimos de intensidad representan el potencial creado por los iones de la red en un sólido cristalino unidimensional. Es importante aclarar que las frecuencias de la luz láser empleadas para confinar los átomos del gas ultrafrío son muy pequeñas en relación a las frecuencias requeridas para el proceso del enfriamiento láser.

Los átomos neutros alcalinos fueron las primeras especies utilizadas en la realización de átomos ultrafríos, sin embargo, átomos con distintas características se incorporaron posteriormente al conjunto de especies que pueden ser atrapados en potenciales de luz. Nos referimos a átomos y moléculas que en forma natural exhiben interacciones de largo alcance ya sea de origen magnético o eléctrico. Dichas interacciones ocurren porque estos átomos y moléculas tienen momentos dipolares magnéticos o eléctricos y por lo tanto se establece entre pares una interacción efectiva de largo alcance [3–7].

Se mencionó varios párrafos atrás que, con la posibilidad experimental de emular del comportamiento de la materia condensada en gases atómicos ultrafríos se desarrollaron otros campos de estudio. Por el lado experimental, la fascinante empresa de crear estados de la materia y estudiar “in situ” sus propiedades, y por otro, el reto teórico de describirlos de primeros principios a través de la teoría de muchos cuerpos. En lo que respecta al contexto experimental, algunos de los desafíos más notables que a la fecha se han logrado incluyen, la realización de juntas de Josephson, la observación de aislantes de Mott, la localización de Anderson, y las fases superfluidas tanto en espacio homogéneo como en redes ópticas, entre otros. Fases más exóticas como la supersólida aún se encuentran como un objetivo a ser alcanzado dentro del contexto de los gases atómicos ultrafríos, pero ya han sido abordadas teóricamente usando como modelo a estos mismos sistemas [8].

El diseño y realización de redes ópticas que confinan átomos o moléculas, y en los que emergen diversas fases como resultado de ambos, interacciones de alcance variable y el número macroscópico de componentes, son una realidad hoy en día. En la figura 5 se muestra una representación esquemática de una bicapa conformada por dos redes con geometría cuadrada, que confinan moléculas fermiónicas [9]. Este tipo de configuración permite elaborar modelos teóricos con los que es posible adentrarse en la descripción de ambas, la superconductividad de alta temperatura, y la superfluidez en ^3He .

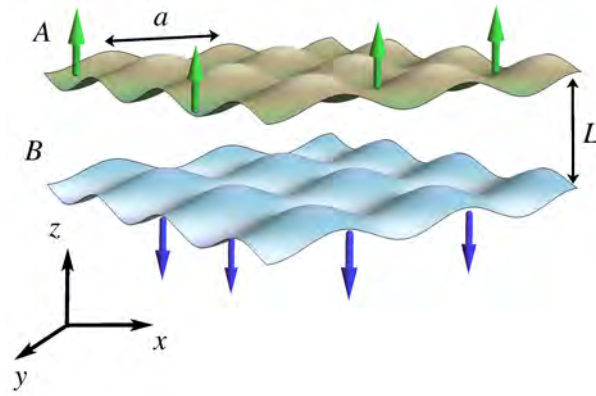


FIG. 5: Representación esquemática de moléculas dipolares de Fermi en un arreglo de bicapa compuesto por dos redes ópticas paralelas en dos dimensiones.

V. CONCLUSIÓN

En estas notas se ha presentado una descripción del fenómeno de superfluidez en redes ópticas. Para dicho propósito primero se discutió el concepto de temperatura y cómo, a través del conocimiento de las leyes que rigen el comportamiento cuántico de la materia, es posible acceder a temperaturas tan cercanas al cero absoluto como 100 nK. Se abordó la descripción de los fenómenos de superfluidez y superconductividad, el primero para átomos de ^4He y gases atómicos ultrafíos alcalinos, y el segundo para el mercurio. En ambos casos se hizo mención de sus características y similitudes. Finalmente, se dedicó la penúltima sección a presentar la forma en la que en un laboratorio son diseñadas y realizadas las redes ópticas en las que ocurre el fenómeno de superfluidez.

Acknowledgments

Este trabajo ha sido apoyado por el proyecto DGAPA-PAPIIT IN108620

-
- [1] K. B. Davis, M. -O. Mewes, M. R. Andrews, N. J. van Druten, D. S. Durfee, D. M. Kurn, and W. Ketterle, *Phys. Rev. Lett.*, **75**, 3969 (1995).
 - [2] J. Bardeen, L.N. Cooper and J. R. Schrieffer, *Phys. Rev.*, **108**, Num. 5, 1175 (1957).
 - [3] A. K. Fedorov, S. I. Matveenko, V. I. Yudson and G. V. Shlyapnikov, *Phys. Rev. A*, **95**, 043615 (2017).
 - [4] A. K. Fedorov, S. I. Matveenko, V. I. Yudson and G. V. Shlyapnikov, *Sci. Rep.*, **6**, 27448 (2016).
 - [5] M. A. Baranov, *Phys. Rep.* **464**, 71 (2008).
 - [6] T. Lahaye, T. Koch, B. Frölich, M. Fattori, J. Metz, A. Griesmaier, S. Giovanazzi, and T. Pfau, *Nature* **448**, 672 (2007).
 - [7] Y. Chen, Z. D. Wang, F. C. Zhang, and C. S. Ting, *Phys. Rev. B* **79**, 054512 (2009).
 - [8] A. Camacho-Guardian and R. Paredes, *Phys. Rev. A* **94** 043638 (2016).
 - [9] G.A. Domínguez-Castro and R. Paredes, *Annalen der Physik*, **20**, 1888316 (2018).

Micromáquinas, pinzas y burbujas

Pedro A. Quinto Su¹

¹Instituto de Ciencias Nucleares, UNAM

November 15, 2021

Abstract

En este documento se muestran algunas de las aplicaciones de haces láser altamente enfocados continuos y pulsados: pinzas ópticas, microexplosiones y motores de vapor. Con la técnica de pinzas ópticas es posible tener un control muy preciso de los objetos pero las fuerzas son del orden de picoNewtons. Las explosiones microscópicas con haces enfocados permiten ejercer fuerzas mucho mayores. Es posible combinar estas dos técnicas en un motor de vapor, donde una partícula microscópica es atraída con las fuerzas ópticas de las pinzas y al acercarse al foco se genera una microexplosión debido al supercalentamiento del agua, repeliendo a la partícula y momentos después el ciclo se repite. Cambiando la estructura de la luz es posible extender las aplicaciones. En particular, cuando la luz se enfoca de manera simultánea en varios puntos se pueden atrapar arreglos de micropartículas o controlar el tipo de flujo con microexplosiones. En el caso de los motores de vapor, las micropartículas se transportan a través de los potenciales generados por los arreglos de focos.

1 Introducción

Los haces láser enfocados de forma pronunciada se han utilizado en numerosas aplicaciones para controlar y manipular objetos microscópicos. En particular con los láseres continuos se pueden atrapar y controlar objetos microscópicos que son atraídos a la región donde la intensidad es máxima. Este tipo de trampas se implementaron por primera vez en la década de 1980 [1] y se conoce como pinzas ópticas.

La Figura 1a muestra un diagrama de una partícula atrapada en el foco del haz. En los ejemplos de este documento se considera que las partículas se encuentran inmersas en líquido, lo que permite evitar fuerzas con alguna superficie del contenedor.

En cambio, cuando se enfocan láseres pulsados con duraciones menores o iguales a decenas de nanosegundos es posible generar explosiones microscópicas en líquido mediante distintos fenómenos físicos como absorción lineal de la luz en un líquido, absorción Bremstrahlung inversa, o absorción multifotónica [2].

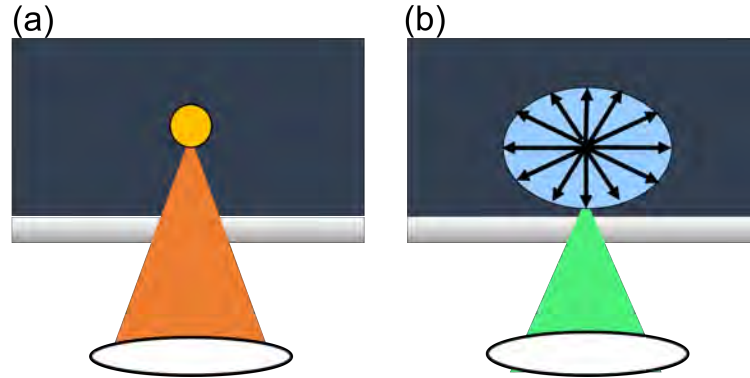


Figure 1: (a) Pinzas ópticas. El láser continuo es enfocado dentro de la muestra que contiene a las micropartículas que son atraídas hacia el punto de máxima intensidad. (b) Microexplosiones. Haz láser pulsado, en el foco se genera una explosión que resulta en la emisión de una onda de choque (no se muestra) y una burbuja que se expande y posteriormente colapsa.

La absorción de la energía en tiempos breves resulta en la emisión de una onda de choque supersónica y una burbuja que se expande y se colapsa en una escala de tiempo que puede ir entre unos cientos de nanosegundos a cientos de microsegundos (Figura 1b). La dinámica de la burbuja se puede utilizar para ejercer fuerzas de muchos órdenes de magnitud mayores a las de las pinzas ópticas con haces continuos. Estos fenómenos se pueden aprovechar para ejercer fuerzas impulsivas de muchos órdenes de magnitud superiores comparadas a las fuerzas que se pueden ejercer con los láseres continuos.

En las siguientes subsecciones se discuten brevemente los siguientes temas: pinzas ópticas, explosiones microscópicas o burbujas, motores microscópicos de vapor, haces enfocados con estructura continuos y pulsados y redes de transporte de micropartículas.

2 Pinzas y burbujas

En esta sección se discuten las aplicaciones básicas para atrapar y generar fuerzas impulsivas a escala microscópica mediante pinzas ópticas y microexplosiones.

Se considera un haz láser con un perfil transversal de intensidad Gaussiano.

2.1 Pinzas Ópticas

La luz láser se enfoca con un objetivo de microscopio con una apertura numérica del orden de 1.2 o mayor.

La fuerza es proporcional al gradiente de la intensidad del haz enfocado,

debido a esto la fuerza es mayor en las direcciones transversales a la dirección de propagación. Las fuerzas tienen el orden de pico Newtons.

$$F \propto \nabla^2 I(r)$$

donde $I(r)$ es la intensidad en función de la posición.

Es importante destacar que las fuerzas son atractivas cuando el índice de refracción de la partícula es superior al del líquido y repulsivas cuando es menor. En el caso de que la intensidad del haz que se enfoca sea Gaussiana ese gradiente resulta en una fuerza restaurativa similar a la de un oscilador armónico para desplazamientos pequeños.

2.2 Explosiones microscópicas

Al enfocar pulsos láser en líquidos se tiene que considerar si el líquido absorbe a la longitud de onda del láser o si es transparente.

En el caso de un líquido transparente es posible que dependiendo de la duración del pulso la luz se absorba por medio de distintos fenómenos físicos generando un plasma que a su vez durante su expansión emite una onda de choque y genera una burbuja que se expande realizando trabajo en el líquido y posteriormente cuando agota su energía se colapsa debido a la presión hidrostática del líquido que la rodea (cavitación).

Cuando el líquido absorbe a la longitud de onda, el medio no se ioniza, pero la absorción rápida de la energía también resulta en la emisión de una onda de choque y una burbuja cavitante.

Utilizando láseres pulsados enfocados se han podido ejercer fuerzas impulsivas a nivel microscópico, por ejemplo para medir propiedades elásticas de glóbulos rojos [3] y nanotubos de carbono [4].

2.3 Motores microscópicos de vapor

Una forma de combinar las fuerzas impulsivas de las explosiones con las pinzas ópticas es utilizar el fenómeno del supercalentamiento del líquido en el cual los objetos microscópicos están inmersos.

El fenómeno de supercalentamiento de un líquido consiste en elevar la temperatura de manera rápida de forma que se excede la temperatura de ebullición sin que el líquido cambie de estado. Este estado supercalentado es metaestable, donde el líquido súbitamente cambia a vapor mediante una explosión.

Para implementar supercalentamiento en una pinza óptica [5], se utilizaron micropartículas que absorben parcialmente a la longitud de onda del láser. Las partículas comienzan abajo del foco y son atraídas hacia este mediante fuerzas ópticas de gradiente. De esta forma al aproximarse al foco de la luz, la temperatura aumenta y calienta el líquido que está en contacto con la micropartícula. En unos instantes parte de ese líquido es supercalentado y unos microsegundos después genera una explosión (Figura 2a) donde la expansión de la burbuja empuja a la partícula abajo del foco y el ciclo se repite. La burbuja desaparece en menos de 5 microsegundos.

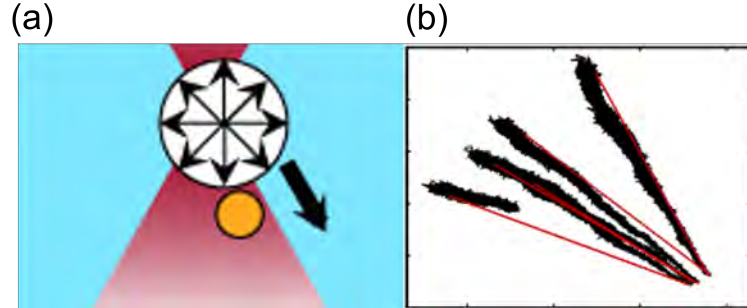


Figure 2: (a) Funcionamiento del motor de vapor. (b) Trayectoria en el plano xy .

El movimiento de la partícula es similar al de un pistón ya que es desplazada por medio de explosiones. La Figura 2b muestra algunos ciclos de la partícula vistos desde arriba (plano xy).

3 Haces enfocados con estructura

En las últimas décadas se han tenido avances que consisten en poder controlar la estructura espacial de la luz láser por medio de elementos ópticos controlados por computadora [6]. Esto se logra por medio de máscaras que pueden imprimir una fase bidimensional a la luz.

3.1 Máscaras de fase bidimensionales

Los elementos ópticos programables más utilizados son los moduladores espaciales de luz y arreglos de microespejos. Consideramos un modulador espacial de luz (SLM por sus siglas en inglés) que es una pantalla de cristal líquido que puede imprimir fases entre cero y 2π en cada uno de sus píxeles. Este dispositivo se controla por medio de imágenes de 8 bits, de manera que se pueden tener 256 valores de fase. La fase que imprime al haz se puede considerar como un factor de transmisión:

$$T = e^{i\phi(x,y)} \quad (1)$$

donde $\phi(x, y)$ es la fase (holograma digital) que se proyecta en el modulador.

Una de las configuraciones más utilizadas consiste en poner el modulador a una distancia igual al foco del lente que enfoca el haz. De esta forma, en el plano focal del lente se tiene un campo que es esencialmente la transformada de Fourier del campo en el modulador.

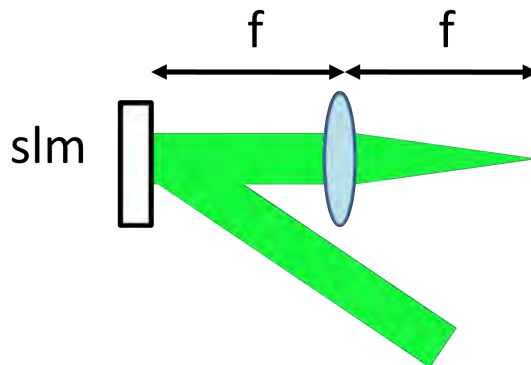


Figure 3: (a) Diagrama simplificado del experimento para generar haces con estructura con un modulador espacial de luz (slm).

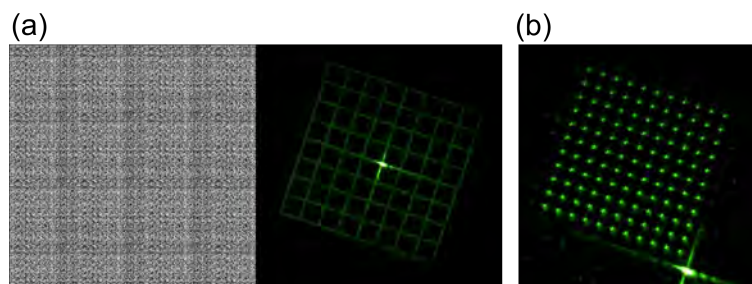


Figure 4: (a) Holograma Digital en el modulador espacial de luz para formar la cuadrícula de la derecha proyectada en el foco del lente. (b) Arreglo de haces enfocados generados por el algoritmo de Gershberg Saxton.

3.2 Arreglos de haces enfocados

Algunas aplicaciones en micromanipulación requieren que la luz se enfoque de forma simultánea en varios puntos. Las fases para generar estos haces se pueden calcular con algoritmos iterativos [7] que utilizan el hecho que el campo enfocado es la transformada de Fourier del campo en el modulador espacial de luz.

3.3 Haces Bessel y Airy

Algunos de los haces que son solución a la ecuación de onda son los haces Bessel que emergen a partir de considerar coordenadas cilíndricas (ρ, z, ϕ) , donde la parte espacial transversal tiene la forma [8]:

(2)

$$A(x, y) = A_m J_m(k_T \rho) e^{im\phi}$$

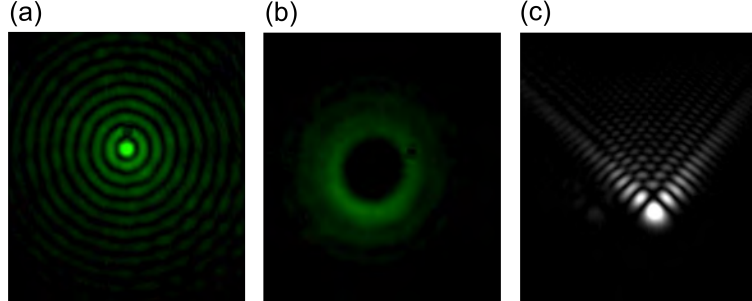


Figure 5: (a) Haz Bessel de orden cero. (b) Vórtice óptico. (c) Haz Airy.

donde la amplitud es una función Bessel de orden m $J_m(k_T\rho)$. El término de fase es proporcional a $e^{im\phi}$ que es una fase azimutal de vórtice porque tiene una singularidad en el centro.

En particular los haces Bessel de orden cero ($m = 0$) se pueden generar con una máscara anular, mientras que sólo utilizando la fase de vórtice se obtienen haces con perfil transversal anular. Algunos ejemplos de estos haces se muestran en la Figura 5a-b.

Otra solución interesante a la ecuación de onda paraxial está dada por los haces Airy [9], que pueden generarse enfocando una amplitud Gaussiana modulada por una fase cúbica resultando en un haz que tiene un perfil transversal de amplitud proporcional a una función Airy (Figura 5c).

Una variante que hemos estudiado de los haces Airy son una versión simétrica [10, 11] donde se toma el valor absoluto de la fase cúbica. Estos haces se pueden utilizar para apilar cientos de micropartículas en estructuras tridimensionales con forma de pirámide [10].

3.4 Haces estructurados continuos

Los láseres continuos con estructura permiten atrapar conjuntos de micropartículas de forma simultánea. La Figura 6a muestra un arreglo cuadrado de 25 trampas y algunas partículas. La Figura 6b muestra un arreglo arbitrario de trampas acomodadas en la forma de las letras 'ICN'.

3.5 Haces estructurados pulsados

Si se utilizan láseres pulsados con máscaras de fase como las que produce un modulador espacial de luz se puede generar una explosión en puntos arbitrarios. La Figura 6c muestra 3 explosiones generadas simultáneamente mediante el enfocamiento de un pulso láser en los vértices de un triángulo. De esta forma es posible controlar la interacción entre las burbujas cnt y ondas de choque [12] para controlar las interacciones con objetos microscópicos.

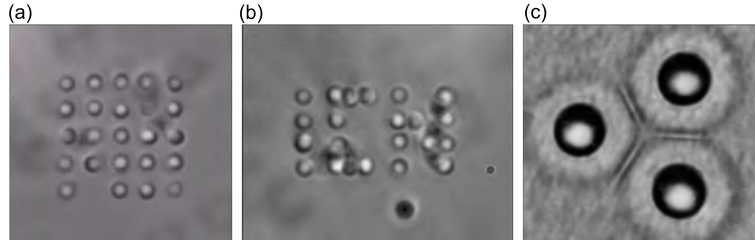


Figure 6: (a) Arreglo de 25 trampas (b) Trampas acomodadas de forma de 'ICN'. (c) Tres microexplosiones simultáneas.

3.6 Redes de transporte con motores microscópicos a través de arreglos de haces estructurados

Para esta aplicación se utilizan arreglos de haces enfocados continuos con una o varias partículas que absorben parcialmente a la longitud de onda del láser como en la aplicación del motor de vapor. Las partículas interactúan con los potenciales ópticos por medio de las fuerzas ópticas de gradiente y las explosiones.

La Figura 7a muestra un dibujo de una red de transporte arbitraria, en la parte superior derecha se muestra la representación de una micropartícula. La fuerza impulsiva ejercida durante una explosión puede empujar a la partícula a la vecindad de otro potencial con el que se generarán nuevos ciclos que pueden inducir el transporte en la red de potenciales. La figura 7b es el resultado de un experimento [13] realizado con trampas colocadas de manera aleatoria y el ancho de las conexiones entre los puntos muestra que conexiones son más frecuentadas en esta red.

Otro efecto interesante es que las explosiones generan una fuerza atractiva en objetos cercanos, por lo que las partículas pueden tener interacciones de larga distancia al transportarse de esta forma.

4 Conclusión

Se mostró que los haces láser enfocados de forma pronunciada constituyen algunas de las mejores herramientas para interactuar con el mundo microscópico. En particular se mostraron los efectos de utilizar láseres continuos y pulsados en aplicaciones como las pinzas ópticas y las burbujas generadas por explosiones. Es posible combinar esas técnicas mediante motores de vapor, donde una partícula microscópica puede interactuar con un potencial óptico continuo a través de fuerzas ópticas y microexplosiones generadas por el supercalentamiento del líquido.

Finalmente, las aplicaciones se pueden extender utilizando luz estructurada generada a partir de máscaras de fase bidimensionales.

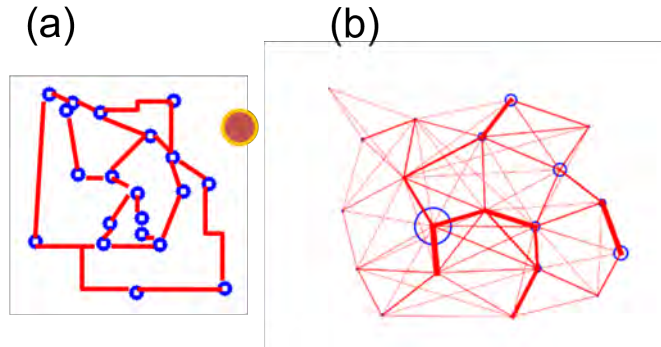


Figure 7: (a) Esquema de una red de transporte. (b) Red de transporte medida con motores microscópicos de vapor.

References

- [1] A. Ashkin, "Acceleration and trapping of particles by radiation pressure", *Phys. Rev. Lett.*, 24, 156-159 (1970).
- [2] Alfred Vogel, Vasan Venugopalan, "Mechanisms of pulsed laser ablation of biological tissues", *Chem Rev*, 103(2):577-644 (2003). doi: 10.1021/cr010379n.
- [3] PA Quinto-Su, C Kuss, PR Preiser, CD Ohl, "Red blood cell rheology using single controlled laser-induced cavitation bubbles", *Lab on a Chip* 11 (4), 672-678 (2011).
- [4] PA Quinto-Su, XH Huang, SR Gonzalez-Avila, T Wu, CD Ohl, "Manipulation and microrheology of carbon nanotubes with laser-induced cavitation bubbles", *Phys. Rev. Lett.* 104 (1), 014501 (2010).
- [5] P.A. Quinto-Su, "A microscopic steam engine implemented in an optical tweezer", *Nature Communications*, 5, 6889 (2014), doi:10.1038/ncomms6889
- [6] D. Grier, "A revolution in optical manipulation", *Nature* 424, 810-816 (2003).
- [7] Gerchberg, R. W.; Saxton, W. O., "A practical algorithm for the determination of the phase from image and diffraction plane pictures", *Optik*. 35: 237-246 (1972).
- [8] Bahaa E. A. Saleh, Malvin Carl Teich, "Fundamentals of Photonics", John Wiley Sons, Inc.
- [9] Siviloglou G. A., Broky J., Dogariu A., Christodoulides D.N., "Observation of Accelerating Airy Beams". *Phys. Rev. Lett.* 99, 213901 (2007).

- [10] P.A. Quinto-Su, and R. Jáuregui, "Optical stacking of microparticles in a pyramidal structure created with a symmetric cubic phase", *Opt. Exp.* 22, Iss 10, pp 12283-12288 (2014).
- [11] R. Jauregui and P.A. Quinto-Su, "On the general properties of symmetric incomplete Airy beams", *JOSA A*, 31, 2484-2488 (2014).
- [12] U.J. Gutiérrez-Hernández, F. De Colle, C.D. Ohl, P.A. Quinto-Su Transient time-delay focusing of shock waves in thin liquids *Journal of Fluid Mechanics* 910, A27 (2021).
- [13] PA Quinto-Su, "Microparticle transport networks with holographic optical tweezers and cavitation bubbles", *Optics letters* 44 (18), 4610-4613 (2019).

La física en el análisis de moléculas y medicamentos biotecnológicos

Dra. Karla Récamier Méndez

Laboratorio de Análisis de Moléculas y Medicamentos Biotecnológicos (LAMMB)

Instituto de Biotecnología UNAM

El Laboratorio de Análisis de Moléculas y Medicamentos Biotecnológicos (LAMMB) perteneciente al Instituto de Biotecnología de la UNAM, es un laboratorio que funciona como un puente entre la UNAM y otras instituciones públicas y laboratorios farmacéuticos de México. Además de realizar investigación y desarrollo, el LAMMB contribuye el registro sanitario de medicamentos biotecnológicos en el país.

Los medicamentos biotecnológicos son proteínas recombinantes tales como factores de crecimiento, hormonas, vacunas, anticuerpos y receptores. Se producen mediante la técnica de ADN recombinante que consiste en expresar un gen de un organismo en OTRO organismo distinto. Para su producción se utilizan bacterias, levaduras, células de insecto y células de mamífero.

Los medicamentos biotecnológicos se utilizan para el tratamiento de una gran variedad de enfermedades. En la tabla 1 se enlistan los medicamentos biotecnológicos innovadores con registro vigente en México. Estos medicamentos suelen tener un precio comercial alto, por lo que resulta de interés desarrollar medicamentos biocomparables. Los medicamentos biocomparables son aquellos que han demostrado ser comparables a los medicamentos innovadores de referencia, una vez que la patente ha expirado. Para demostrar que un medicamento biotecnológico es biocomparable, se realiza el mayor número de estudios posibles para demostrar su calidad, eficacia y seguridad. Los estudios a realizar están establecidos en la Farmacopea de los Estados Unidos Mexicanos: Métodos generales de análisis (MGA) y Monografías para cada medicamento, las Guías ICH: *International Conference for Harmonization*, la Farmacopea de los Estados Unidos de América y la Farmacopea Europea, entre otros.

En términos de la ley general de Salud (Art. 376) un registro sanitario es una Autorización Sanitaria con la cual deberán contar los medicamentos, dispositivos médicos, plaguicidas y nutrientes vegetales. Sólo podrá ser otorgado por la Secretaría de Salud y ésta ejercerá las atribuciones de regulación, control y fomento sanitario, a través de la Comisión Federal para la Protección contra Riesgos Sanitarios (COFEPRIS) en lo relativo a la ley general de la salud (Art. 17 bis).

Tabla 1

Medicamentos biotecnológicos innovadores con registro vigente en México

#	Denominación Genérica	Denominación Distintiva	Titular	Indicación terapéutica
1	Eritropoyetina theta	EPORATIO	Lemery, S.A. de C.V.	Anemia sintomática.
2	Idarizumab	PRAXBIND	Boehringer Ingelheim Promeco, S.A. de C.V.	Reversión del efecto anticoagulante de
3	Insulina glargina	TOUJEO	Sanofi-Aventis Deutschland GmbH	Diabetes mellitus tipo 1 y tipo 2
4	Insulina humana recombinante	NOVOLIN R	Novo Nordisk A/S	Para el tratamiento de diabetes mellitus.
5	Insulina humana recombinante isófana	NOVOLIN N	Novo Nordisk A/S	Para el tratamiento de la diabetes mellitus.
6	Insulina humana recombinante isófana /	NOVOLIN 70/30	Novo Nordisk A/S	Diabetes mellitus tipo I.
7	Interferón beta 1b	EXTAVIA	Novartis Pharma AG	Esclerosis múltiple remitente recurrente y
8	Ixequizumab	TALTZ	Eli Lilly and Company	Psoriasis
9	Mepolizumab	NUCALA	GlaxoSmithKline México, S.A. de C.V.	Asma
10	Molgramostim	GRAMAL	Probiomed, S.A. de C.V.	Factor estimulante de colonias de granulocitos.
11	Necitumumab	PORTRAZZA	Eli Lilly and Company	Cáncer pulmonar
12	Nivolumab	OPDIVO	Bristol-Myers Squibb Holdings Pharma Ltd.	Cáncer pulmonar
13	Nonacog gamma	RIXUBIS	Baxalta US Inc.	Hemofilia B
14	Octocog alfa	RECOMBINATE	Baxter AG	Hemofilia A.
15	Pembrolizumab	KEYTRUDA	Schering-Plough, S.A. de C.V.	Melanoma y cáncer de pulmón de células
16	Secukinumab	COSENTYX	Novartis Pharma AG	Psoriasis, espondilitis anquilosante activa y
17	Somatropina	NORDITROPIN	Novo Nordisk A/S	Disfunción de la hormona de crecimiento.
18	Somatropina	SAIZEN	Merck, S.A. de C.V.	Disfunción de la hormona de crecimiento.
19	Tocilizumab	ROACTEMRA SC	F. Hoffman-La Roche Ltd.	Artritis reumatoide
20	Trastuzumab	HERCEPTIN SC 600	F. Hoffman-La Roche Ltd.	Cáncer de mama
21	Turoctocog alfa	NOVOEIGHT	Novo Nordisk A/S	Hemofilia A.

Obtener el registro sanitario para un medicamento biotecnológico es un proceso que consta de varias etapas y depende de si es innovador o biocomparable.

- 1) Investigación y desarrollo: 10-15 años.
- 2) Producción: Buenas prácticas de fabricación (GMP por sus siglas en inglés *Good Manufacturing Practices*): 5 años.
- 3) Control de calidad: Rama analítica y rama documental.
- 4) Ensayos preclínicos: 2 años.
- 5) Ensayos clínicos: 2 años.
- 6) En paralelo: estudios de estabilidad.

Toda la información obtenida desde la investigación-desarrollo hasta los ensayos clínicos y estabilidades se juntan en un *dossier* y se somete a análisis por la autoridad regulatoria. Por lo tanto, se opera bajo un Sistema de Gestión de Calidad que se ajusta a las normativas vigentes.

En LAMMB se realizan pruebas de calidad, de caracterización fisicoquímica y actividad biológica. Algunas de estas son: determinación de endotoxinas, ADN residual, proteína hospedera, pH, osmolalidad, apariencia, mapeo peptídico, cromatografía de exclusión molecular, cromatografía de intercambio iónico, espectrometría de masas, electroforesis capilar y sulfidrilos libres.

Asimismo, cabe destacar la participación del LAMMB en la colaboración internacional con la Coalición de Innovación en Preparación ante Epidemias (CEPI por sus siglas en inglés). Este proyecto forma parte de una iniciativa para la evaluación de vacunas candidatas contra la COVID-19, además de revisar las pruebas y ensayos clínicos frente a variantes de preocupación del SARS-CoV-2.

A pesar de que estas pruebas son principalmente químicas o biológicas, los equipos utilizados para realizarlas operan bajo varios principios de Física. A continuación algunos ejemplos:

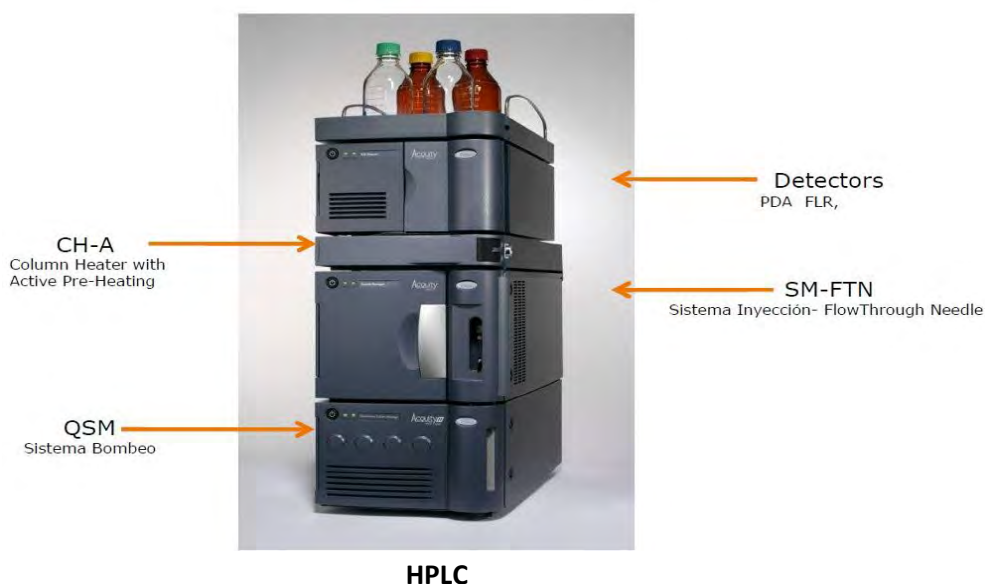
- 1) Cromatografía líquida de alta resolución (HPLC por sus siglas en inglés *High Performance Liquid Chromatography*). Es una técnica cromatográfica que permite la separación y análisis de diferentes mezclas de compuestos como fármacos, proteínas y compuestos químicos. Se basa en la separación de solutos entre dos fases, una fase móvil que está en movimiento constante con respecto a la fase estacionaria. El sistema puede ser configurado con los detectores de TUV, PDA, FLR y QDa:

***Detector TUV:** detector de luz ultravioleta / visible (UV/Vis), opera en longitudes de onda de 190 a 700 nm.

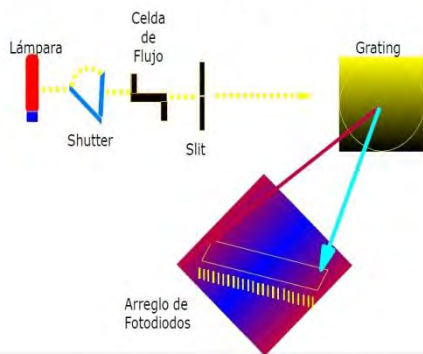
***Detector óptico PDA** (detector de arreglo de fotodiodos): es un espectrómetro de luz UV/Vis que funciona entre 190 y 500 nm.

***Detector FLR:** detector de fluorescencia que tiene un rango de longitud de onda de excitación de 200 a 890 nm, un rango de longitud de onda de emisión de 210 a 900 nm, admite velocidades de datos de hasta 80 Hz y ofrece capacidad de escaneo 3D para facilitar el desarrollo de métodos.

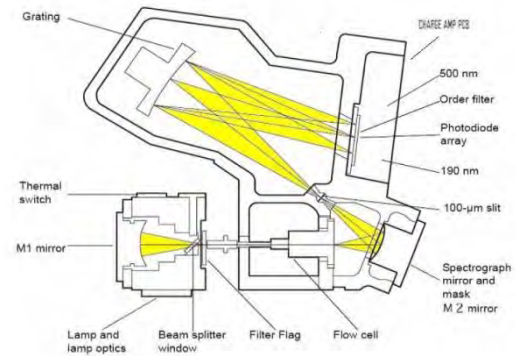
***Detector QDa:** detector de masas cuadrupolo sencillo. La espectrometría de masas se basa en la razón masa/carga (m/z). Se pueden analizar biomoléculas con pesos moleculares de hasta 40 kDa.



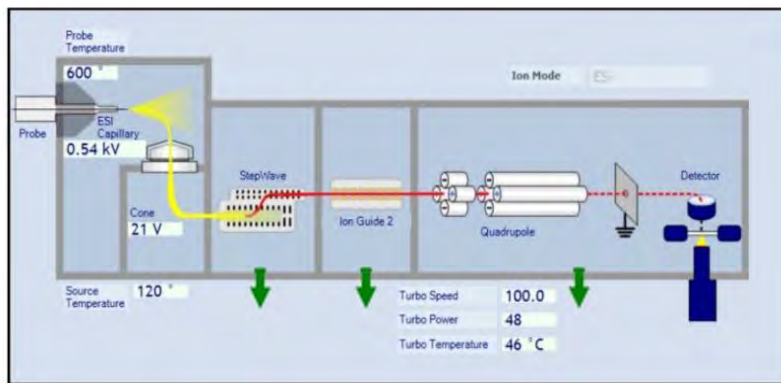
La Óptica del Arreglo de Diodos



Trayectoria de la luz en la Óptica del Acquity UPLC PDA Detector



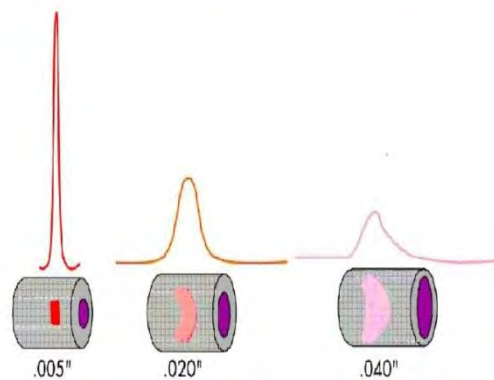
Detector PDA



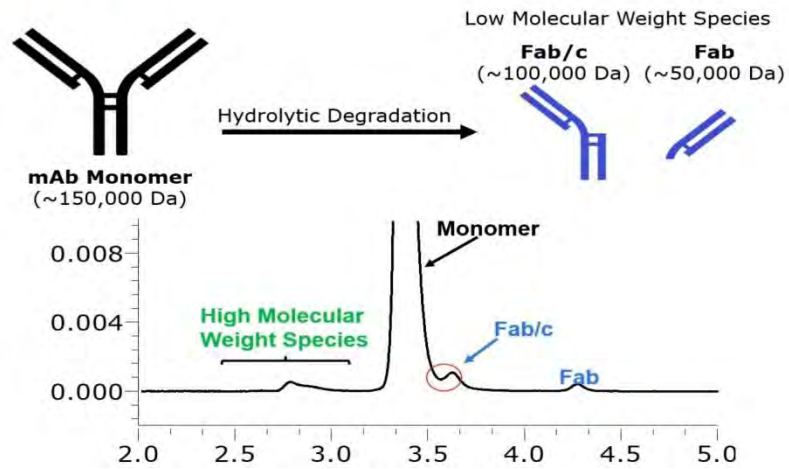
Detector QDa

Para los análisis es necesario seleccionar un rango de longitud de onda que incluya todas las bandas de absorción de los analitos. Además, hay un efecto de la tubería en la difusión de los componentes. Se pueden tener presiones de hasta 15000 psi, logrando una baja dispersión de los analitos, una máxima relación señal/ruido y una mayor resolución.

Efecto de la tubería en la difusión de los componentes



mAb Centric SEC

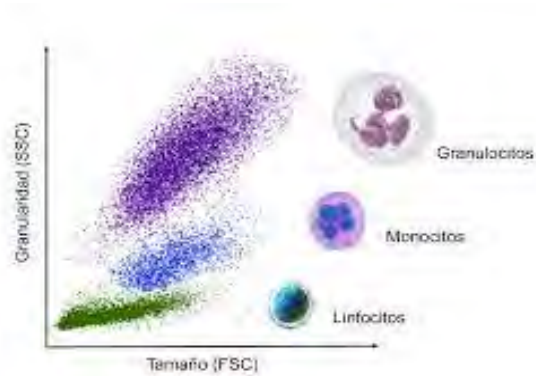


Perfil obtenido en cromatografía de exclusión molecular

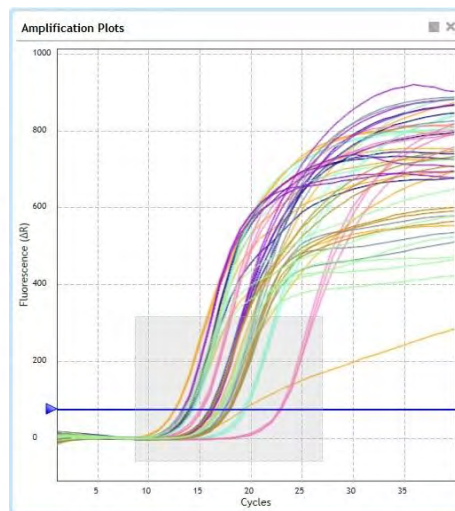
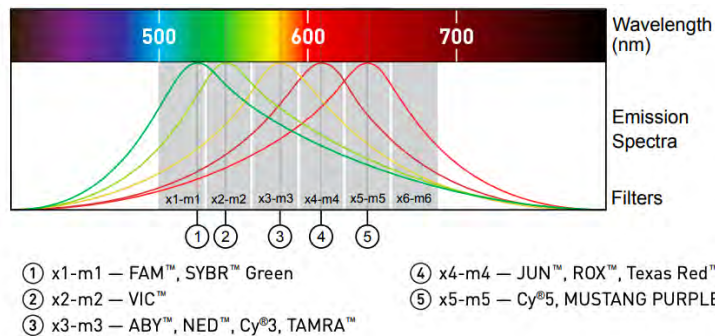
- 2) Electroforesis capilar: Técnica de separación basada en la diferente velocidad de migración de las especies cargadas bajo la acción de un campo eléctrico. Bajo la acción de alto voltaje (1-30 kV) los componentes de la muestra migran de manera diferencial por el capilar. A medida que pasan por una ventana del capilar, el detector mide la absorbancia y transmite la señal a la computadora. Es una técnica de alta eficiencia de separación de moléculas por medio de diferencias en su relación carga/masa. Se utiliza para determinar la pureza de los fármacos.



- 3) Citometría de flujo: Medir características ópticas y de fluorescencia de una célula o cualquier partícula.



- 4) qPCR: Reacción en cadena de la polimerasa cuantitativa o en tiempo real. Utilizado para amplificar, detectar y analizar ácidos nucleicos. El diseño del sistema combina un termociclador, un sistema óptico avanzado con una fuente de excitación LED y un software de análisis de datos. El equipo admite hasta 6 filtros ópticos (azul-FAM, verde-VIC, amarillo-ABY, naranja-JUN, rojo-Cy[®]5 y rojo intenso-Cy[®]5.5).



- 5) Espectrómetro: Absorbancia 200-1000 nm; Fluorescencia: 250-850 nm. Las proteínas tienen un espectro de absorción ultravioleta característico alrededor de 280 nm, predominantemente debido a sus aminoácidos aromáticos como tirosina y triptófano. Por lo tanto, las proteínas y los péptidos absorben luz UV en proporción con su contenido de aminoácidos aromáticos y la concentración total. Si se conoce el coeficiente de extinción molar de la proteína, entonces se puede utilizar la ley Beer-Lambert para cuantificar de manera precisa la cantidad de proteína por absorbancia UV.

Donde:

a_m : coeficiente de extinción molar

c : concentración del analito

l : longitud del paso óptico en cm



De tal forma, podemos apreciar como la Física es una Ciencia que puede llegar a englobar aspectos de interés muy variados, incluso en otras ramas de conocimiento.

Física a la nanoescala

Penélope Rodríguez^{*} Lauro Oliver Paz[†] and Andrés Botello[‡]

Instituto de Física, Universidad Nacional Autónoma de México

Keywords: first keyword, second keyword, third keyword

I. INTRODUCCIÓN

El desarrollo de la mecánica cuántica en la primera mitad del siglo XX ha contribuido de manera excepcional al conocimiento de la física de la materia en general, y a la física del estado sólido en particular. En el transcurso del siglo pasado, el entendimiento de la relación entre composición y estructura de los sólidos por una parte, y sus propiedades por otra, ha llevado a la invención de una gran variedad de dispositivos que se pueden caracterizar como de estado sólido. En las últimas décadas, la investigación en la física de estado sólido ha desvelado conexiones entre las propiedades estructurales, elásticas, electrónicas, magnéticas y ópticas de los sólidos y ha resultado en el descubrimiento de fenómenos fundamentales como los efectos Hall cuánticos, superconductores de alta temperatura crítica, el efecto de magnetoresistencia gigante y láseres de estado sólido, por mencionar algunos.

Además, el avance en el crecimiento controlado de materiales, que permite el dominio de las propiedades de la materia con base en su composición y estructura, junto con la invención de técnicas experimentales de caracterización, ha sentado las bases para el florecimiento de los sistemas de baja dimensionalidad y las nanoestructuras, que presentan a su vez nueva física y nuevas perspectivas tecnológicas. Estas notas tienen como objetivo proporcionar al lector un panorama general de la física a escalas reducidas - *la física a la nanoescala*.

El punto de partida es la física del estado sólido o de la materia condensada, ya que la física de los sólidos cristalinos es más fácil de describir mediante teorías aproximadas que explotan la simetría traslacional inherente a los cristales. Con base en estas teorías, que fueron formuladas para cristales tridimensionales por la obvia razón de que es el estado en las que se encuentran naturalmente, se pueden hacer predicciones de los efectos de disminuir la dimensionalidad. ¿Qué diferencias hay entre las propiedades de los cristales en 3d, 2d y 1d?

Antes de entrar en materia, vale la pena preguntarse por qué es que estamos interesados en estudiar la física a la nanoescala. Daremos nuestra visión en la primera sección de estas notas. Posteriormente, haremos una concisa introducción a la física del estado sólido con énfasis en la estructura electrónica y en algunas propiedades y en los cambios que hay cuando se disminuyen las dimensiones del sistema. Como veremos, el problema es complejo y se requieren muchas aproximaciones para llegar a obtener

* penelope.rz@fisica.unam.mx

† oliver_paz@fisica.unam.mx

‡ botello@fisica.unam.mx

resultados útiles. Aquí resulta importante resaltar que en nuestros días, el avance en la comprensión de la física a la nanoescala depende de cálculos numéricos que nos permitan acceder a cantidades observables, ya sea para interpretar y o explicar los experimentos, o para predecir sus resultados. Esta sinergia entre teoría-simulación y experimentos ha sido fundamental para el rápido avance del área en las últimas décadas. Así pues, la tercer sección tendrá como objetivo introducir brevemente los conceptos detrás del método computacional más utilizado para la simulación de las propiedades de los materiales: la teoría del funcional de la densidad.

Con esto tendremos las bases para poder entrar al mundo de la física a la nanoescala. Las siguientes secciones, entonces, tendrán como objetivo dar una *probadita* del área a través de ejemplos puntuales y representativos de cómo funciona esta sinergia.

II. NANOCIENCIAS

Formada por el prefijo *nano* que indica la mil-millonésima parte de algo, la palabra nanociencia se refiere al estudio de los fenómenos que ocurren cuando las escalas espaciales son del orden de nanómetros ($1\text{nm} = 1 \times 10^{-9}$ metros). En particular, los nanomateriales son aquellos que tienen al menos una dimensión espacial menor a 100 nm *y* que muestran propiedades distintas a las de sus análogos macroscópicos.

Debido a que muchos proceso químicos y biológicos ocurren a escalas de nanómetros, las nanociencias son inherentemente una rama del conocimiento interdisciplinaria. Sin embargo, en el área de la física química, se hace énfasis en utilizar los conceptos y teorías físicas para entender los fenómenos que ocurren a dichas escalas. Es a eso a lo que nos referimos por *física a la nanoescala*. Como ya adelantamos, el punto de partida será la física del estado sólido, a la cual haremos una breve introducción a continuación.

III. FÍSICA DEL ESTADO SÓLIDO

La física está llena de analogías, y antes de entrar de lleno con la física del estado sólido, vale la pena introducir dos conceptos fundamentales que son la base de dichas analogías. El primer

concepto es el del gas (homogéneo) de electrones libres. Aquí, los electrones no están sujetos a ningún potencial. En el límite de un gas muy diluido, uno puede pensar que la dinámica de cada electrón está determinada sólo por su energía cinética que es igual a $E = p^2/2m$ y a éste le llamamos el gas de electrones no interactuantes. El segundo concepto es el de cuasi-partícula. En el modelo estándar, el número de partículas está fijo y sus propiedades también. En los sólidos se forman excitaciones que se comportan *como si* fueran partículas en el sentido de que pueden transferir momento. Ejemplos son las vibraciones en el cristal, el concepto de agujero o los electrones que se mueven con masas efectivas distintas a las de los electrones libres.

En mecánica cuántica, el estado fundamental de un sistema es aquel que posee mínima energía. Normalmente los sistemas atómicos se encuentran en este estado, por lo que su estudio es de vital importancia para un gran número de ramas de la ciencia.

La descripción precisa de las propiedades de un sistema - sea un ensamble de átomos como lo puede ser una molécula aislada, o bien, los átomos que definen un cristal - la obtenemos, según la mecánica cuántica, gracias a la ecuación de Schrödinger: $\mathcal{H}E = \Psi E$, donde el Hamiltoniano \mathcal{H} está dado por:

$$\mathcal{H} = \mathcal{H}_{nucl.} + \mathcal{H}_{el} + \mathcal{H}_{el-nucl.} \quad (1)$$

donde

$$\mathcal{H}_{nucl.} = \overbrace{-\sum_{I=1}^{N_z} \frac{\hbar^2}{2M_I} \nabla_{\mathbf{R}_I}^2}^{cinetica} + \overbrace{\frac{1}{8\pi\epsilon_0} \sum_{\substack{I,J=1 \\ I \neq J}}^{N_z} \frac{Z_I Z_J e^2}{|\mathbf{R}_I - \mathbf{R}_J|}}^{potencial(Coulombica)} \quad (2)$$

$$\mathcal{H}_{el} = -\sum_{i=1}^{N_e} \frac{\hbar^2}{2m_i} \nabla_{\mathbf{r}_i}^2 + \frac{1}{8\pi\epsilon_0} \sum_{\substack{i,j=1 \\ i \neq j}}^{N_e} \frac{e^2}{|\mathbf{r}_i - \mathbf{r}_j|} \quad (3)$$

$$\mathcal{H}_{el-nucl} = -\frac{1}{4\pi\epsilon_0} \sum_{I=1}^{N_z} \sum_{j=1}^{N_e} \frac{Z_I e^2}{|\mathbf{R}_I - \mathbf{r}_j|} \quad (4)$$

Si conociéramos Ψ , entonces tendríamos la información necesaria para describir las propiedades del estado base de cualquier sistema, como los antes mencionados. Sin embargo, estamos en general interesados en casos de sistemas con múltiples electrones los cuales interactúan con múltiples núcleos, lo cual hace la situación un poco más complicada. Esto es debido a que ahora enfrentamos un

problema de N_e electrones y N_z núcleos, con $N_e, N_z \sim 10^{23}$. Debido al alto nivel de complejidad, resulta imposible resolver esta ecuación en la práctica, por lo que se hacen algunas aproximaciones. Las más importantes serán:

Aproximación adiabática o de Born-Oppenheimer:

- Núcleos 10^4 a 10^5 veces más pesados que los electrones
- dinámicas más lentas
- energía cinética de los núcleos puede ser despreciada
- separación de los problemas: núcleos y electrones

Por simplicidad, en este punto se introducen las unidades atómicas. De esta manera, el Hamiltoniano con posiciones nucleares fijas se ve como:

$$\hat{H} = \underbrace{-\frac{1}{2} \sum_i \nabla_{\mathbf{r}_i}^2}_{\hat{T}} - \underbrace{\sum_{i,j} \frac{Z_i}{|\mathbf{R}_i - \mathbf{r}_j|}}_{\hat{V}_{ext}} + \underbrace{\sum_{i \neq j} \frac{1}{|\mathbf{r}_i - \mathbf{r}_j|}}_{\hat{U}} + \underbrace{\sum_{i \neq j} \frac{Z_i Z_j}{|\mathbf{R}_i - \mathbf{R}_j|}}_{\hat{V}_{NN}} \quad (5)$$

Este Hamiltoniano corresponde al movimiento electrónico en la presencia de un potencial externo producido por los núcleos atómicos. Es importante mencionar que la repulsión nuclear (\hat{V}_{NN}) contribuye a la energía total únicamente por un factor constante, por lo que remover dicho término de la ecuación no afectará a la función de onda correspondiente.

$$\hat{H} = \underbrace{-\frac{1}{2} \sum_i \nabla_{\mathbf{r}_i}^2}_{\hat{T}} - \underbrace{\sum_{i,j} \frac{Z_i}{|\mathbf{R}_i - \mathbf{r}_j|}}_{\hat{V}_{ext}} + \underbrace{\sum_{i \neq j} \frac{1}{|\mathbf{r}_i - \mathbf{r}_j|}}_{\hat{U}} \quad (6)$$

A la ecuación anterior se le conoce como Hamiltoniano electrónico, donde \hat{T} representa la energía cinética de los electrones, \hat{U} la repulsión electrónica y \hat{V}_{ext} la interacción entre núcleos y electrones.

Cristales:

Los cristales son una idealización matemática de los sólidos, en donde se supone simetría traslacional *ad infinitum*. En este sentido, un cristal está determinado a partir de una red o retícula



Figure 1. Un cristal está compuesto de una retícula infinita definida por vectores de traslación y una base atómica.

definida por un conjunto de vectores de traslación. Si se define un punto cualquiera de la retícula como el origen, los otros puntos de la retícula se pueden alcanzar a través de combinaciones lineales de los vectores unitarios de la red: $T_{\vec{n}} = n_1\vec{a}_1 + n_2\vec{a}_2 + n_3\vec{a}_3$, donde $\vec{n} = (n_1, n_2, n_3)$ es un vector de números enteros y $\vec{a}_1, \vec{a}_2, \vec{a}_3$ son los vectores unitarios de la red.

Además de la retícula, existe una base atómica, es decir una serie de coordenadas que definen las posiciones de los átomos con respecto a cada punto de la retícula. La retícula y la base atómica definen el cristal. Se define como celda unitaria el elemento que se repite a lo largo de todo el espacio y que define el cristal (ver figura 1).

A. Teorema de Bloch

El carácter traslacional de los cristales permite proponer una solución a la función de onda con el supuesto de que ésta debe siempre puede ser compuesta de una onda plana y una función con la misma periodicidad del cristal.

$$\psi_{\mathbf{k}}(\mathbf{r}) = e^{i\mathbf{k}\cdot\mathbf{r}}u_{\mathbf{k}}(\mathbf{r}) \quad (7)$$

Los índices que definen la onda plana están relacionados con el cristal y definen el llamado espacio recíproco, cuyos vectores unitarios están relacionados por $\mathbf{A}\mathbf{B} = 2\pi\mathbf{I}$, donde \mathbf{A} es una matriz 3×3 (para el caso 3D) donde las columnas son los vectores unitarios de la red, \mathbf{B} es una matriz 3×3 , donde los renglones son los vectores de la red recíproca y \mathbf{I} es la matriz identidad 3×3 . El equivalente a la celda unitaria (de Wigner Seitz) en el espacio recíproco es la zona de Brillouin.

Nota que mientras la función de onda en el espacio real en dos puntos equivalentes por simetría traslacional difiere por una fase, en el espacio recíproco la función de onda tiene la periodicidad de la red recíproca. Por esta razón, los vectores de onda \mathbf{k} , se usan para describir la estructura electrónica de los sólidos cristalinos.

B. Estructura electrónica

Para determinar las propiedades electrónicas de los materiales, se calcula su *estructura electrónica*.

La estructura de bandas no da el valor de la energía para cada estado electrónico en función de su momento cristalino \mathbf{k} . Debido a que \mathbf{k} es, en general, un vector tridimensional, no hay forma de visualizar $E_{\mathbf{k}}$ en función de \mathbf{k} (necesitaríamos una gráfica de 4 dimensiones!!).

La forma de representar $E_{\mathbf{k}}$, es a través de un diagrama en el que el eje x es una coordenada a lo largo de un camino que recorre la Zona de Brillouin pasando por los puntos de alta simetría. En la figura de abajo, un camino posible sería seguir la línea roja.

La estructura de bandas nos dice en qué direcciones el cristal tiene estados disponibles. La pendiente de cada una de las curvas de dispersión está relacionada con la velocidad de grupo de los electrones en esa dirección.

En el ejemplo anterior, podemos ver que cerca de la energía de Fermi (que separa los estados ocupados de los estados desocupados), hay estados en un algunas de las direcciones, mientras que no hay estados en otras direcciones.

Esto quiere decir que si sometemos la muestra a una diferencia de potencial finita ($\neq 0$), siempre habrá electrones disponibles para generar una corriente. En otras palabras, el sistema conduce electricidad. Los electrones que conducen tendrá un momento cristalino que corresponde a la dirección \mathbf{k} y una velocidad de grupo proporcional a la pendiente de la curva de dispersión.

Si en cambio no hubiera ningún estado disponible al rededor de nivel de energía de Fermi, entonces el sistema no tendría electrones para generar una corriente hasta que la diferencia de potencial fuera mayor a la brecha de energía entre estados ocupados y desocupados. Ese es el caso de un semiconductor.

Otra forma de visualizar lo arriba descrito es a través de la densidad de estados. En este caso

se integran todos los estados electrónicos y se obtiene una curva más sencilla que sólo nos dice en cada energía si cuantos estados electrónicos hay disponibles sin tener información sobre su momento cristalino o su velocidad de grupo.

Aún si esta forma de representar la estructura electrónica contiene menos información, es útil ya que no necesariamente el camino que se utiliza (líneas rojas en la figura arriba) contiene la información más relevante del sistema. Por ejemplo, podría haber un sistema que a lo largo de una trayectoria que pasa por los puntos de simetría parece tener una brecha de energía (ser un semiconductor) y que sin embargo la densidad de estados nos dice que sí hay estados en el nivel de Fermi, es sólo que estos estados no están en ningún camino de alta simetría.

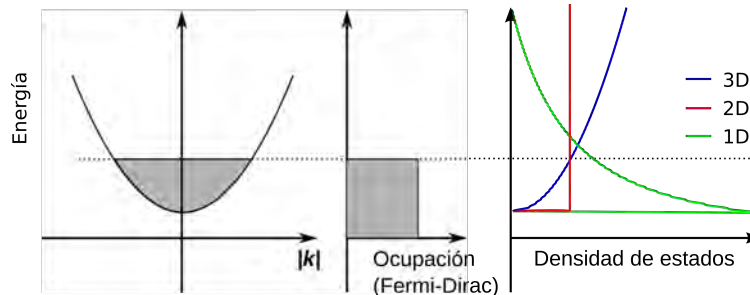


Figure 2. Izq. Dispersión electrónica (energía contra momento cristalino) de un gas de electrones. Centro. Función de distribución de ocupaciones. Der. Densidad de estados para sistemas con dispersión cuadrática en una (1D) dos (2D) y tres dimensiones (3D).

Modelos de los estados electrónicos:

- semiclásicos
- electrones de valencia
- electrones en campos promedio

IV. TEORÍA DEL FUNCIONAL DE LA DENSIDAD

Aún con las aproximaciones realizadas, la función de onda depende de $4N$ variables, tres espaciales y una de espín para cada uno de los N electrones del sistema. Surge entonces de manera natural la

pregunta: ¿ hay otra manera de reemplazar la complicada función de onda de N electrones, por una metodología más sencilla de tratar computacionalmente?

Si, uno de los métodos más exitosos para reducir el problema de la dinámica de los electrones en la materia y en los sólidos cristalinos en particular se basa en la teoría del funcional de la densidad, o DFT por sus siglas en inglés. Es una teoría de campo medio basada en el método variacional detrás de las técnicas de Hartree y Hartree-Fock, pero donde ahora se introduce la densidad electrónica siguiendo los pasos de uno de los primeros métodos que emplean la densidad electrónica como Thomas-Fermi.

En un sistema, la densidad electrónica se define como el número de electrones por unidad de volumen en un estado dado:

$$\rho(\mathbf{r}) = N \int \cdots \int |\Psi(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N)|^2 ds_1 d\mathbf{x}_2 \dots d\mathbf{x}_N \quad (8)$$

de modo que al integrar la densidad electrónica, obtenemos el número total de electrones de nuestro sistema:

$$\int \rho(\mathbf{r}) d\mathbf{r}_1 = N \quad (9)$$

A. Teoremas fundamentales de DFT

La Teoría del Funcional de la Densidad se originó en 1964, cuando Hohenberg y Kohn publicaron un importante artículo en *Physical Review*. Los teoremas probados en este artículo representan la base teórica sobre la que se han construido todas las teorías modernas del funcional de la densidad. El primer teorema de Hohenberg-Kohn nos menciona que:

- a) El potencial externo $v(\mathbf{r})$ es un funcional de la densidad electrónica $\rho(\mathbf{r})$.

Este teorema nos habla también que no solo el potencial externo es un funcional único de la densidad, sino también la función de onda del estado fundamental (base):

$$\Psi_0(\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_N) = \Psi[\rho_0(\mathbf{r})] \quad (10)$$

así como el valor esperado de cualquier observable del estado fundamental, dándonos a la vez la

energía y la densidad del estado base.

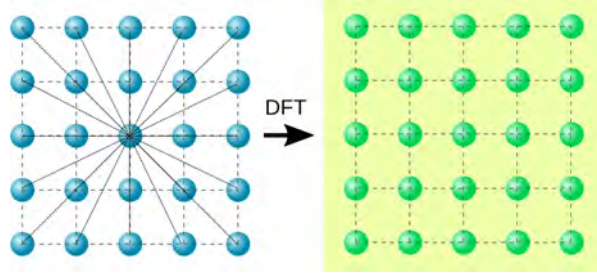


Figure 3. La teoría del funcional de la densidad (DFT) bajo el ansatz de Kohn y Sham mapea un problema de muchos cuerpos interactuantes a un problema de cuasi-partículas que se comportan *como si fueran electrones* independiente moviéndose en un potencial efectivo (a determinar) y que reproducen la densidad electrónica del problema original interactuante.

El segundo teorema de Hohenberg-Kohn se encarga de esto, al definir una propiedad importante de este funcional.

b) La densidad electrónica que minimiza la energía sobre todo el funcional es la densidad electrónica real, que corresponde a la solución completa de la ecuación de Schrödinger.

Del principio variacional, la energía del estado fundamental se obtiene como:

$$E_{v,0} = \min_{\Psi'} \langle \Psi' | \hat{H} | \Psi' \rangle \quad (11)$$

B. Ecuaciones de Kohn-Sham

Kohn y Sham desarrollaron la metodología de DFT al sistema de N electrones no interactuantes en un potencial externo, en las llamadas ecuaciones de Kohn-Sham. Si expresamos todos los términos del Hamiltoniano, en términos de la densidad electrónica $\rho(\mathbf{r})$, tenemos:

$$E[\rho(\mathbf{r})] = T_s[\rho(\mathbf{r})] + J[\rho(\mathbf{r})] + E_{Ne}[\rho(\mathbf{r})] + E_{ex}[\rho(\mathbf{r})] \quad (12)$$

donde $T_s[\rho(\mathbf{r})]$ es la energía cinética del sistema de partículas interactuantes, $J[\rho(\mathbf{r})]$ es el término de Coulomb, $E_{Ne}[\rho(\mathbf{r})]$ es la interacción núcleo-electrón, y finalmente $E_{ex}[\rho(\mathbf{r})]$ es el término de

correlación-intercambio, que nos describe como interactúan los electrones entre sí.

Aplicando el método variacional, se pueden derivar las ecuaciones de Kohn-Sham, las cuales se tienen que resolver de manera numérica a modo autoconsistente:

$$\left[\frac{-\hbar^2}{2m} \nabla_{\mathbf{r}}^2 + \hat{V}^{eff}(\mathbf{r}, \rho(\mathbf{r})) \right] \Phi_i(\mathbf{r}) = \epsilon_i \Phi_i(\mathbf{r}) \quad (13)$$

con

$$\hat{V}^{eff} = \hat{V}_{el-ion} + \int d\mathbf{r}' \frac{\rho(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} + \frac{\delta E_{xc}[\rho(\mathbf{r})]}{\delta \rho(\mathbf{r})} \quad (14)$$

Así, gracias al método variacional, se puede, a partir de un conjunto de funciones de onda de prueba $\phi_i^0(\mathbf{r})$, se puede resolver el problema de forma iterativa:

1. A partir de ϕ_i^0 , construir $\rho(\mathbf{r})$ usando

$$\rho^0(\mathbf{r}) = \sum_i |\phi_i^0(\mathbf{r})|^2 \quad (15)$$

2. Resolver la ecuaciones de Kohn-Sham usando ρ^0
3. Obtener el nuevo conjunto de funciones de onda ϕ_i
4. Calcular la densidad ρ
5. Comparar la densidad ρ^0 y ρ . Si la diferencia está dentro de un criterio de tolerancia (arbitrario), esa es la densidad del estado base y $E_{tot} = \sum_i \epsilon_i$ es su energía. De lo contrario, es necesario mezclar ρ_0 y ρ para obtener una nueva ρ_0 y volver al paso 1 hasta obtener *autoconsistencia*

De acuerdo con su formulación, la DFT es una teoría de *energías totales*. Así, en su formulación más sencilla, nos permite calcular energías relativas, por ejemplo, energías de cohesión o formación.

V. CLUSTERS METÁLICOS: DEFINICIONES Y APLICACIONES

Hablando de sistemas específicos a la nanoescala, uno que resalta en especial son Los cúmulos metálicos (o **clusters**). Estos se pueden definir como son agregados desde unos cuantos átomos hasta

millones de ellos. Pueden estar conformados desde un solo tipo de átomo hasta 2 o más especies de ellos [1]. Su estructura geométrica depende fundamentalmente de su tamaño dado: esta pueda variar desde estructuras amorfas de baja simetría, hasta arreglos un poco más simétricos (e.g. icosaedros, decahedros, etc..) a estructuras cristalinas cuando estos clusters alcanzan tamaños nanométricos (e.g arreglo FCC). Los clusters pueden ser estudiados - tanto experimental como computacionalmente - a modo general en el **estado gas** (libres, sin ningún tipo de interacciones con otros sistemas), o bien **soportados en superficies** (i.e. óxidos metálicos como MgO, TiO₂, CeO₂, por ejemplo), o bien, **pasivados por ligando orgánicos** (i.e. aminoácidos con un grupo tiol), ver Figura 4.

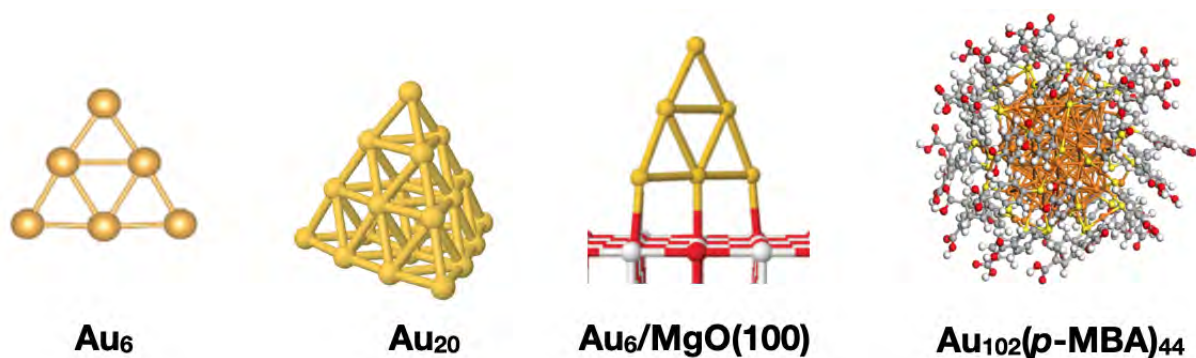


Figure 4. Modelos computacionales a partir de cálculos y simulaciones DFT de clusters metálicos de oro (Au), tanto en: estado gas Au₆ y Au₂₀, soportados como Au₆/MgO(100) así como pasivados Au₁₀₂(p-MBA)₄₄. Imágenes tomadas de las referencias [2-5].

Una de las características esenciales de los cúmulos o clusters metálicos es que al ser tan pequeños, tienen una alta relación entre el número de átomos que ocupan la superficie vs. los átomos que ocupan posiciones centrales. En pocas palabras, esto implica que los clusters metálicos tienen un gran número de átomos ocupando sitios superficiales[6]. Este tipo de características físicas los hace ideales para ser empleados como nanocatalizadores - es decir, compuestos inorgánicos que aceleran una reacción química dada - ya que muchas de estas reacciones se dan en los sitios superficiales de estos compuestos metálicos, ver Figura 5.

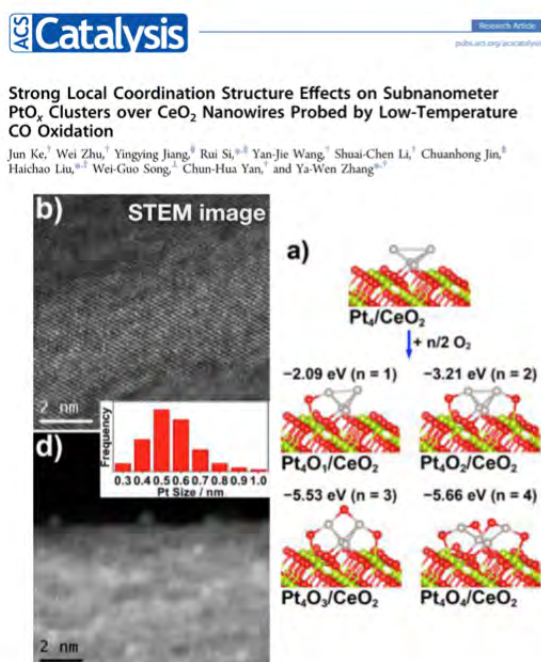


Figure 5. Aplicaciones de clusters sub-nanométricos de platino (Pt) soportados, empleados para la reacción química de oxidación de monóxido de carbono (CO) en dióxido de carbono (CO₂). Los modelos computacionales son generados a partir de extensos cálculos y simulaciones DFT, mientras que las imágenes experimentales fueron tomadas mediante microscopía electrónica (HAADF-STEM y STM) [7, 8].

Debido que estamos trabajando en el mundo de la nanoescala, es importante resaltar que las propiedades fisico-químicas de los clusters serán diferentes a aquellas de los metales en el bulto, debido principalmente al confinamiento cuántico. En el bulto, cada metal tiene llenas (o parcialmente llenas) sus bandas electrónicas. En el bulto, existe un continuo de niveles energéticos y la ocupación electrónica de las bandas permitidas da lugar a: **aislantes**, con una gran brecha de energías prohibidas entre bandas llenas y vacías, de modo que no hay conducción de electrones; **semiconductores** (con una brecha energética comparable a la energía térmica, de modo que puede permitirse el efecto de conducción de electrones), y **metales** (con una banda de conducción parcialmente llena, de modo que los electrones se desplazan libremente). Estas propiedades cambian drásticamente en función del tamaño, y a ello se le conoce como *transición metal-aislante*, o por sus

siglas en ingles *metal-insulator transition (MIT)* [9]; ver Figura 6.

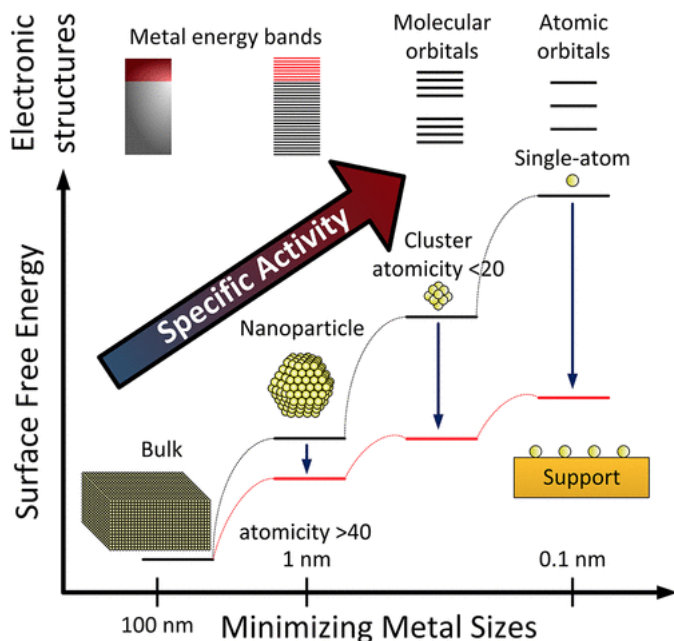


Figure 6. (arriba) Esquema donde se muestra la estructura electrónica de un metal en el bulto - con sus correspondientes bandas energeticas, y la transición hacia un sistema cada vez más pequeño y finito (cluster metálico), donde ya los niveles energéticos se han discretizado y el sistema se comporta más como una molécula o bien punto cuántico. (abajo). Imágen tomada de referencia [10].

En si, es difícil determinar cuando un cluster es *pequeño*, *mediano*, o bien *grande*. Diferentes descripciones pueden ser empleadas, sin embargo, podemos decir que los *clusters pequeños* son aquellos en donde sus propiedades fisico-químicas cambian considerablemente de acuerdo a su tamaño atómico (*ca.* $< 1\text{nm}$). Esto implica que pequeños cambios en el número total de átomos que conforman al cluster tiene implicaciones muy grandes en su estructura geométrica y por ende en sus propiedades. En otras palabras, cada átomo que se agregue cuenta!. Para clusters *medianos*, con cada nuevo átomo que se agregue, las propiedades fisico-químicas en general empiezan a cambiar de una manera más suave, aunque pueden existir ciertas discontinuidades o cambios bruscos. Finalmente, para clusters *grandes*, ya sus propiedades empiezan a parecerse al metal en el bulto, y ya empiezan a ensamblarse

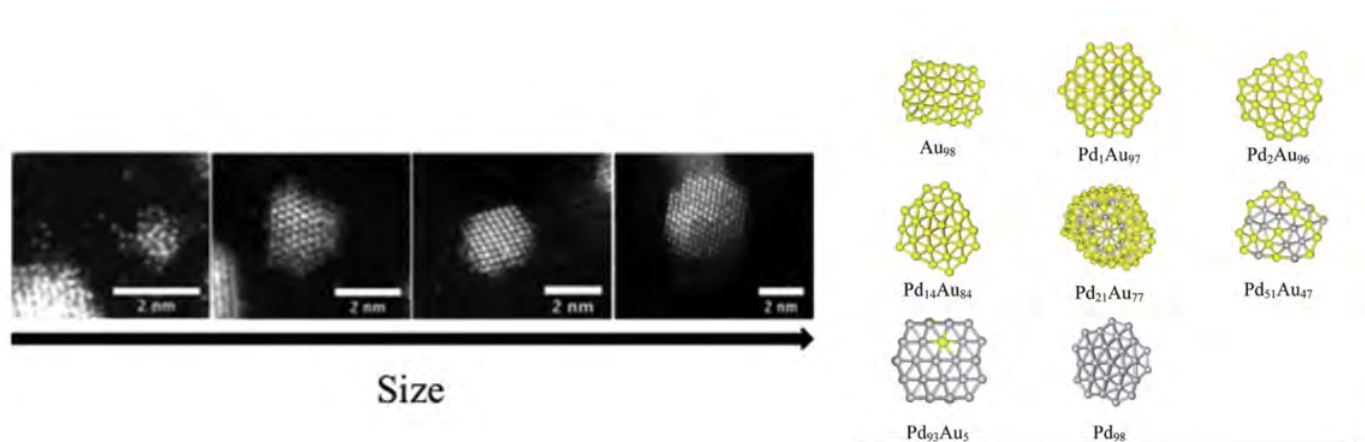


Figure 7. Estructuras más estables de clusters nanometricos bimetalicos de Pd-Au de tamaño de 98 atomos encontradas por un Algoritmo genético (GA) acoplado a un potencial semi-empirico de Gupta, y su comparación con imagenes experimentales de microscopia electronica HAADF-STEM. Imagenes tomadas de la referencia [11].

en configuraciones tridimensionales, adoptando sobre todo estructuras cada vez mas cristalinas.

A. Búsqueda de configuraciones de clusters estables

Una vez que entramos a las diversas líneas de investigación que involucran el uso de nanoestructuras así como clusters metálicos, nos topamos con un pequeño pero muy importante dilema a la hora de querer reconciliar resultados experimentales con modelos computacionales: **¿Cuales son las configuraciones estables de estos clusters, a un tamaño dado?**. Además, la palabra *estable* puede implicar muchas situaciones, particularmente relacionadas a condiciones de temperatura bajo las cuales se da el experimento, de presión de gases (*e.g.* ambiente oxidativo con O₂, o reductivo con H₂), si esta interaccionando con ligandos organicos, etc, por ejemplo.

Es por ello que en los últimos años, ha habido un auge en el desarrollo de técnicas matemáticas de optimización enfocadas a encontrar las estructuras más estables (*e.g.* correspondientes a las de

mínima energía) tanto de nanopartículas como de clusters metálicos (o bimetálicos, involucrando 2 metales para formar una nanoaleación). Entre ellas destacan metodologías basadas en Algoritmos Genéticos (GA)[12], Basin Hopping Monte Carlo[13], Simulated-Annealing[14], Particle Swarm Optimization[15], entre otras. La pregunta clave es, *¿cómo se lleva a cabo una optimización global de clusters metálicos?*. De manera general, implica acoplar la metodología de optimización de nuestra preferencia a un modelo físico que nos describa correctamente las interacciones metal-metal que se dan en el sistema. Para ello se pueden realizar aproximaciones basadas en potenciales semi-empíricos, o bien, al nivel de cálculos mecánico-cuánticos. En términos generales, realizar una optimización global implica hacer una búsqueda exhaustiva de posibles configuraciones estables de la geometría - así el arreglo químico más favorable que se da entre 2 metales en el cluster o nanopartícula - dado el nivel de teoría empleado.

Dadas las limitaciones encontradas en el pasado en términos de acceso a avanzados centros de supercomputo para realizar este tipo de búsquedas, una de las primeras aproximaciones teóricas fue emplear potenciales intermoleculares. Estos en sí dependen de una construcción matemática que relaciona una parte atractiva y otra repulsiva. Un clásico ejemplo de ello es el potencial Lennard-Jones, el cual ha sido empleado en la descripción de sistemas del tipo gases nobles[16]. Sin embargo, un potencial del tipo LJ no es apropiado en la descripción de sistemas metálicos, por ejemplo metales de transición (Au, Pt, Ag, Pd, entre otros). Es por ello que se desarrollaron los potenciales de muchos cuerpos semi-empíricos; es decir, que el potencial toma en cuenta a todos los átomos sintiendo la interacción entre ellos, mientras que ciertos parámetros implícitos del potencial dependen de valores experimentales como constantes de red, etc. A estos potenciales se le conoce del tipo Gupta [17]. Este tipo de potenciales ha sido empleado en la búsqueda de las configuraciones de mínima energía de clusters bimetálicos de Au-Pt a tamaño de 98 átomos, ver Figura 7. Para encontrar las configuraciones estructurales de mínima energía, así como del arreglo químico preferencial entre los dos metales se empleó un Algoritmo Genético en todas las composiciones químicas[11].

Gracias al desarrollo de infraestructura de cómputo de alto rendimiento, dentro de la UNAM contamos con nuestra propia computadora para realizar cálculos en paralelo masivos, esto dentro



Figure 8. Supercomputa "Miztli" UNAM. Actualmente cuenta con 430 nodos con procesadores Intel Xeon E5, para un total de 8040 cores de computo intensivo.

de Supercomputo UNAM (Miztli), ver Figura 8. Esto nos ha permitido implementar algoritmos de optimización global basados Basin Hopping Monte Carlo (BH-DFT), acoplado al software de cálculos mecanico-cuanticos DFT periodico, Quantum Espresso; ver Figura 9. Gracias a ello, hemos podido entender el crecimiento de clusters subnanometricos de platino (Pt), así como platino-cobre (Pt-Cu) soportados en la superficie del oxido de cerio $\text{CeO}_2(111)$ [18, 19]. Además, hemos estudiado las propiedades fisico-quimicas de dichos sistemas, sobre todo en lo relacionado a su reactividad y posible actividad catalitica en base a sus geometrias y estructura electronica.

En nuestra implementación, el código de optimización BH-DFT realiza búsquedas sistematicas para configuraciones estables, empleando "*kicks*", y *atomic-swaps*; mientras que el código de DFT periodico Quantum Espresso se encarga de minimizar las fuerzas que sienten los átomos metálicos del cluster y los lleva a un mínimo local. Para ilustrar como funciona el algoritmo, la rutina de optimización empieza por generar un átomo metálico, digamos Pt, en algún lugar aleatorio de la supercelda asignada. El siguiente átomo se agrega a una distancia del primer átomo de acuerdo a su enlace metal-metal, usualmente el orden de 1.50 - 2.50 Å. Átomos subsecuentes, son posicionados a una distancia similar, con sus nuevas coordenadas aleatorias *xyz*. Esto empieza a generar una configuración tridimensional, ya sea en el estado gas, o posicionada a una cierta altura de la superficie del oxido donde se quieren estudiar estos clusters metálicos; ver Figura 9. Esta estructura es entonces *relajada* por el software DFT Quantum Espresso, y a la estructura obtenida se le aplica una distorsión via un "*kicks*". Esto distorciona aleatoriamente todas las coordenadas *xyz* de nuestro

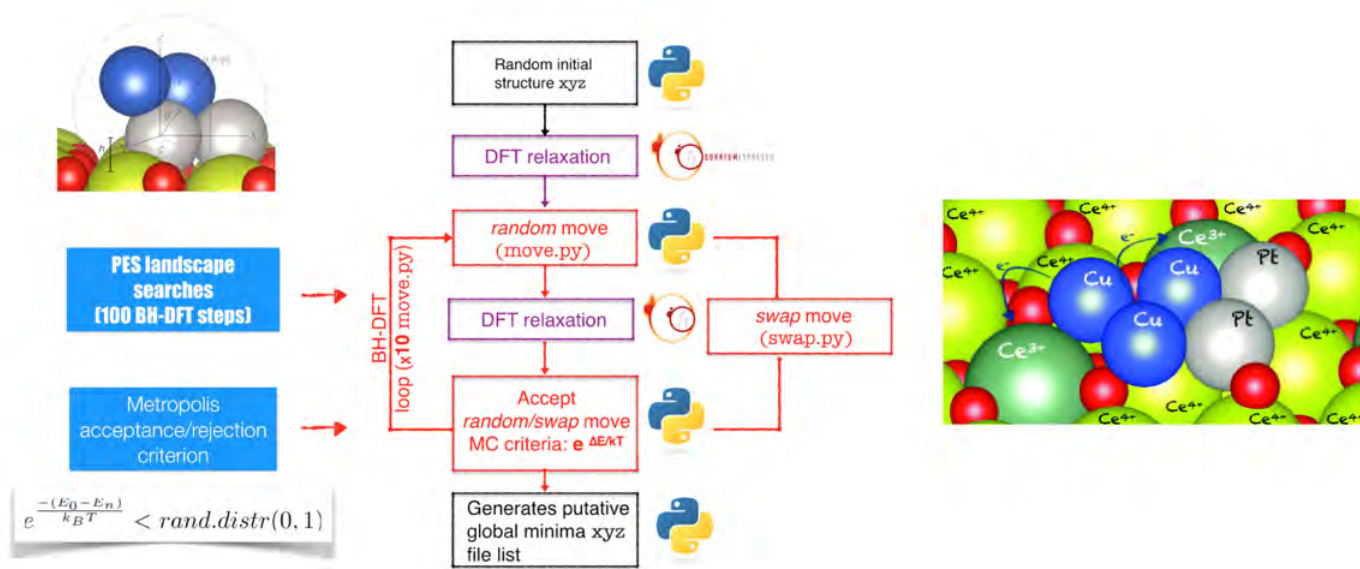


Figure 9. (izquierda) Diagrama de nuestra implementación para el algoritmo de optimización global Basin Hopping Monte Carlo (BH-DFT), acoplado al software de cálculos mecanico-cuanticos DFT periodico, Quantum Espresso. (derecha) Configuración de mínima energía encontrada por nuestro algoritmo BH-DFT para el caso del cluster bimetalico Pt₂Cu₃ soportado en la superficie CeO₂(111).

cluster metalico, lo cual implica que al relajarlo de nueva cuenta, terminemos en una estructura de mas baja energia (mas estable). Sin embargo, aqui entra el criterio Monte Carlo de nuestro algoritmo ($e^{-\Delta E/kT}$). Este termino basicamente acepta a todas las configuraciones de mas baja energia que las anteriores; sin embargo, asigna una cierta probabilidad kT de que algunas estructuras inestables puedan ser aceptadas. Esto se hace para lograr una variedad en la distribución de estructuras, y que eventualmente nos permita explorar vastas areas configuraciones y llegar al minimo global en el menor tiempo posible. Finalmente, en nuestro grupo tambien hemos implementado otros tipos de algoritmos de busqueda de optimizacion global basada en movimientos aleatorios ("random-search"), particularmente cuando buscamos los sitios de adsorcion de moleculas organicas del tipo cisteina sobre nanoparticulas metalicas del tipo oro (Au) [20].

VI. FRENTE EXPERIMENTAL

Un número creciente de investigaciones en ciencia a la nanoescala se realizan conjugando estudios teóricos y experimentales que se llevan a cabo de manera consecutiva o en paralelo. Las herramientas teóricas nos permiten generar una predicción y a la vez una confirmación de resultados experimentales, mientras que el enfrentamiento de ambos quehaceres mejora sustancialmente la interpretación fenomenológica.

En la actualidad existe todavía una brecha entre los resultados experimentales y las simulaciones teóricas en muchos campos, y en algunas áreas de las nanociencias es en dicha brecha donde radica la complejidad de los nano-sistemas. Por ello, ambos frentes suman esfuerzos para acortar la brecha: por su lado, los experimentos avanzan hacia un mayor control de variables, incluyendo resolución atómica, tanto en cuestión de síntesis como en caracterización, mientras que los cálculos teóricos incrementan su capacidad para lograr representar fidedignamente sistemas con cada vez más grados de complejidad disminuyendo el costo computacional.

En el área de investigación de nanocúmulos metálicos encontramos un claro ejemplo de lo anterior. El progreso en técnicas experimentales de síntesis como las fuentes de nanocúmulos por medio de pulverización catódica y selección de masas por tiempo de vuelo, la separación por tamaños por cromatografía líquida de alta resolución (HPLC, por sus siglas en inglés) y la cristalografía con rayos X (XRC, por sus siglas en inglés), nos ha permitido producir nanocúmulos con un número definido de átomos (recordemos que en la nanoescala el aumento o disminución de tan sólo un par de átomos puede cambiar las propiedades ópticas, geométricas, catalíticas, etc. de un sistema), y caracterizarlos con resolución atómica. Lo anterior se extiende a nanocúmulos con ligandos moleculares en donde también se cuenta con control del número exacto de moléculas. De esta manera uno de los objetivos es la obtención de sistemas experimentales controlados con un número mínimo de átomos (tanto los que conforman los nanocúmulos metálicos como los que conforman las moléculas ligando), tal que sus características puedan ser reproducidas en cálculos teóricos reduciendo el número de aproximaciones [21–25].

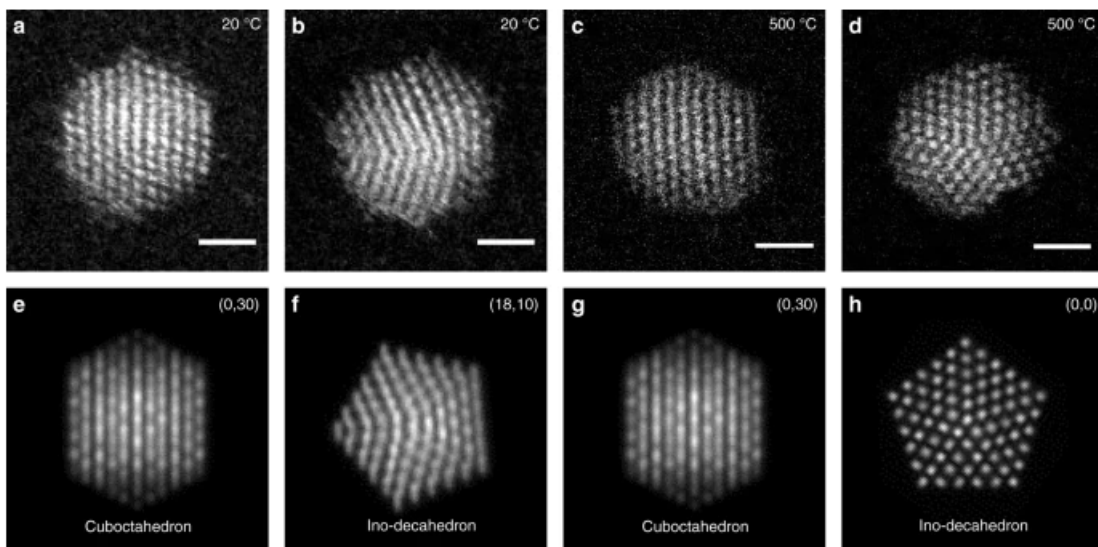


Figure 10. a–d) Imágenes de HAADF STEM de cúmulos de oro Au561; e–h) Simulaciones equivalentes de dispersión de electrones en multi-capa del cuboctaedro e ico-decahedro a diferentes orientaciones. [26]

A continuación se describirán brevemente algunas técnicas experimentales para la síntesis y caracterización de nanocúmulos metálicos con y sin ligandos moleculares.

A. Síntesis de Nanocúmulos

1. Fuente de Nanocúmulos

La generación de cúmulos con tamaño seleccionado consiste en tres etapas: producción de cúmulos, enfoque y aceleración de haz de cúmulos y selección de masa. Cada una de ellas está asociada a una sección de la fuente de cúmulos como se muestra en la figura 11.

En la primera sección existe una cámara de condensación en la que, gracias a la aplicación de un alto voltaje de radiofrecuencia, plasma de argón es encendido cerca de un blanco del metal con el

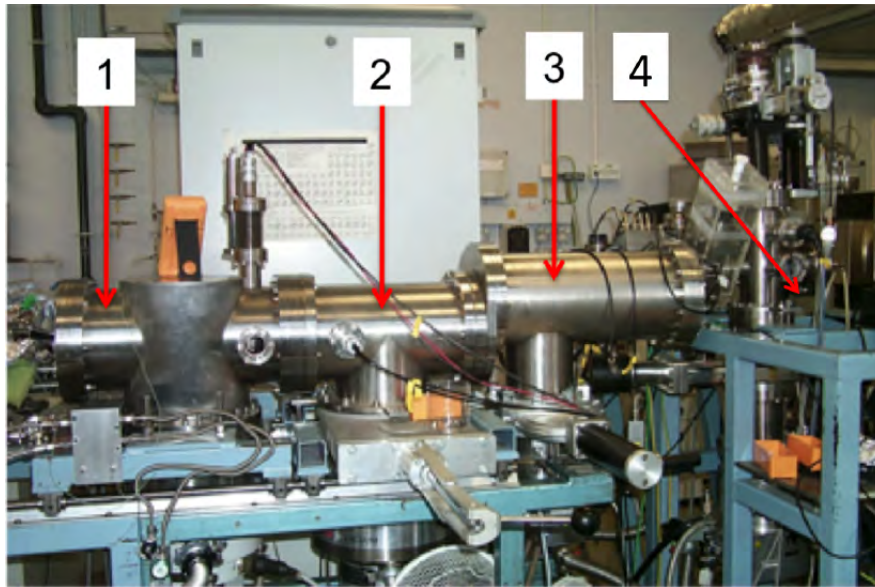


Figure 11. Fuente de cúmulos. Las secciones correspondientes a la pulverización catódica y la formación de cúmulos (1), enfoque y aceleración de haz de cúmulos (2), selección de masa (3) y deposición (4) están indicadas en la imagen. (Fotografía de la fuente de cúmulos G12 del Laboratorio de Investigaciones en Física de la Nanoescala, Universidad de Birmingham, Reino Unido.)

que se deseen formar los nanocúmulos. Como consecuencia, el blanco será negativamente cargado y generará un voltaje negativo (self-bias), induciendo así un campo eléctrico que ocasiona pulverización catódica de átomos metálicos calientes debido al bombardeo de los átomos de argón. Los átomos son consecutivamente condensados en cúmulos metálicos con una amplia distribución de tamaños ($2 - 10^5$ átomos cada uno) por la acción de un gas de helio inyectado con una presión típica de 0.2 mbar.

En la segunda sección existe un conjunto de lentes einzel con un potencial negativo que enfoca el haz ionizado. En el centro de los lentes einzel existe un conjunto de platos de deflexión que ajustan la trayectoria del haz para ingresar a la tercera sección por la acción de un campo eléctrico en dirección transversal que dispersa el haz en un plano horizontal de acuerdo a la masa de los cúmulos.

Cuando el haz entra en la tercera sección, se le aplica un pulso de aceleración producido por el aumento de potencial momentáneo de la placa inferior la misma que está conectada a un fuente de alto voltaje. Con esta aceleración todos los iones ganan el mismo momento y por lo tanto todos los iones con la misma masa adquirirán la misma velocidad dirigida hacia la parte superior de la cámara. Cuando todos los iones de la misma masa se encuentran en la región intermedia de la cámara se evita que suban más por medio de un segundo pulso con la misma amplitud pero esta vez proveniente de la placa superior, la cual también se encuentra conectada a un alto voltaje. De esta forma, los iones con masa diferente son transmitidos en función de la diferencia de tiempo entre los pulsos: el tiempo que le toma a los iones viajar de la primera a la tercera región varía, de manera que el segundo pulso puede ser aplicado tal que se posicione el haz con masa seleccionada (de entre los haces paralelos con masas específicas) a la altura requerida de la apertura al final de la cámara. Este principio hace que la masa de los iones transmitidos sea independiente de la energía del haz. La resolución de la masa está determinada por la razón del ancho de la apertura de salida y el desplazamiento lateral.

Al final de la tercera sección, un haz de cúmulos con masa seleccionada es enfocado en una cámara de alto vacío con una presión de 10^{-8} mbar para ser depositado (Fig.12) [21, 27].

B. Caracterización de Nanocúmulos

1. Microscopía de Fuerza Atómica

El comportamiento de dos átomos sin enlace pero en proximidad cercana como función de la separación internuclear r puede ser descrita por el potencial de Lennard-Jones (Fig. 13):

$$V(r) = (A/r^{12}) - (C/r^6) \quad (16)$$

Donde A y C son constantes asociadas a cada tipo de átomo que involucra la profundidad de la pared de potencial y la distancia a la cual el potencial inter-partícula es cero.

Como función de la distancia, la fuerza interatómica va de fuerza altamente repulsiva, causada por interacciones coulombianas y representada por el primer término en la ecuación (15), a fuerza

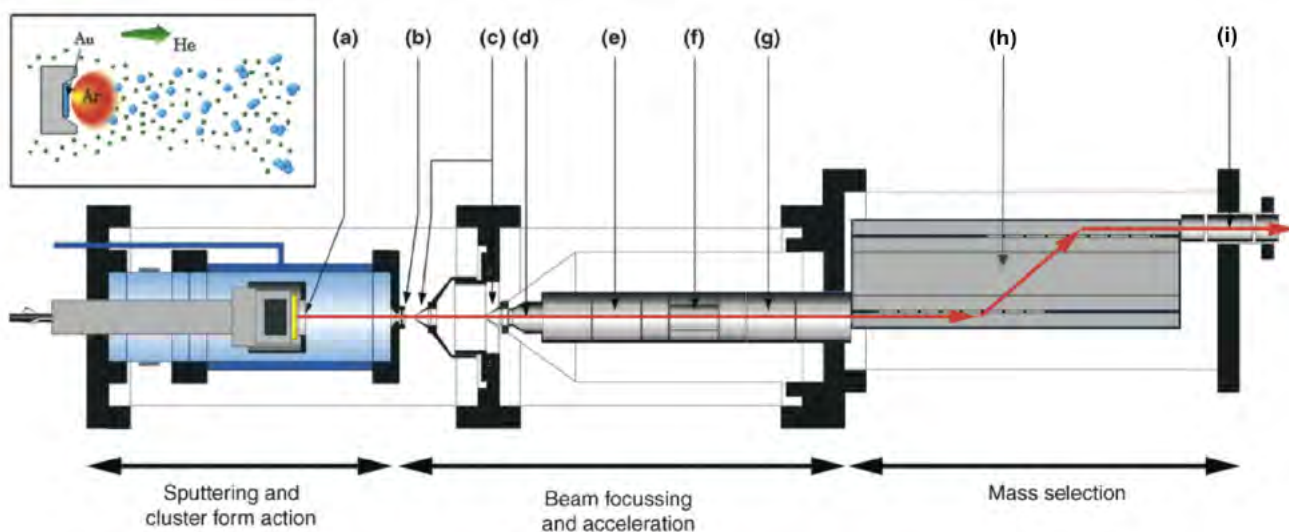


Figure 12. Diagrama de la fuente de cúmulos con selección de masa por tiempo de vuelo. (a) Pistola de magnetrón con blanco metálico; (b) boquilla de diámetro ajustable; (c) electrostatic skimmers; (d) lente de alto voltaje para extracción; (e) primer lente einzel; (f) platos de deflexión; (g) segundo lente einzel; (h) región de filtro de masas por tiempo de vuelo (TOF) (i) tercer lente einzel.

atractiva, debido a fuerzas de van der Waals, representadas por el segundo término de la misma ecuación, pasando por cero a una distancia normalmente de $2A$. Antes de cruzar ese punto, los átomos están todavía en contacto, a partir de ese punto, ya en el régimen de fuerza atractiva, los átomos primero se encuentran en contacto intermitente y luego de estar en un punto de equilibrio, pasan a una región en la que el contacto se pierde por completo.

La microscopía de fuerza atómica está basada en las fuerzas electromagnéticas atractivo-repulsivas entre los átomos de la superficie de la muestra y los átomos de la punta del voladizo del microscopio. La interacción de fuerza provoca que el voladizo sea deflactado según la ley de Hooke. Conforme la punta escanea la muestra, variaciones en la topografía de la superficie ocasionan cambios de la interacción de fuerzas y por lo tanto el desplazamiento angular del voladizo.

Los microscopios de fuerza atómica pueden ser utilizados en dos modos principales de operación

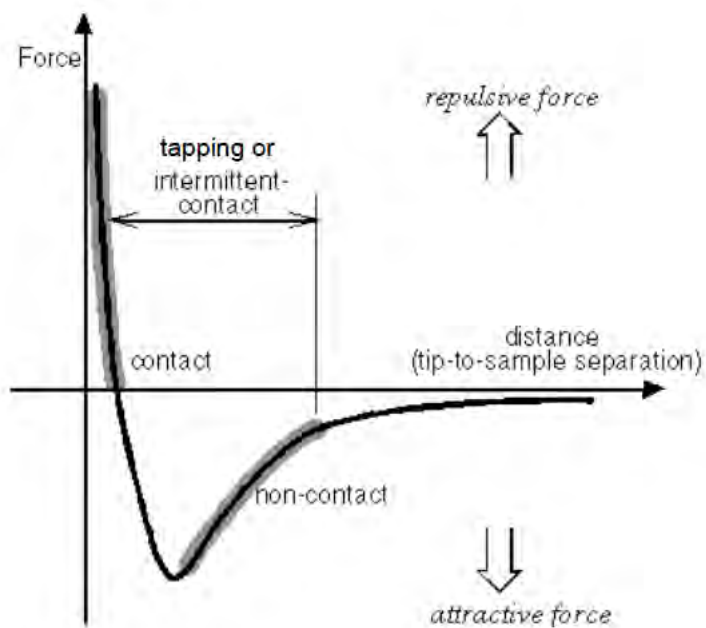


Figure 13. Curva de fuerza interatómica contra distancia. La gráfica indica las regiones correspondientes a los modos de contacto, contacto intermitente y no contacto de la microscopía de fuerza atómica. En el régimen de contacto el voladizo se encuentra a unos pocos angstroms de la superficie de la muestra, donde la fuerza interatómica es repulsiva. La región de contacto intermitente empieza cuando la punta se separa de la superficie aproximadamente el equivalente a la longitud de un enlace químico. Cuando el espacio entre punta y muestra es del orden de décimos de cientos de angstroms la punta alcanza el régimen de no contacto.

distintos: modo de contacto, modo de no contacto.

En el modo de contacto (CM-AFM, por sus siglas en inglés) la proximidad de la punta a la muestra cae en el régimen en el que la fuerza repulsiva es dominante. En este rango la separación punta-muestra no es compresible, lo que resulta en el doblamiento del voladizo. La fuerza intrínseca del voladizo depende de la constante de resorte y la deflexión a la que está sujeto. La fuerza se puede calcular partiendo de la ley de Hooke que describe un resorte elástico lineal:

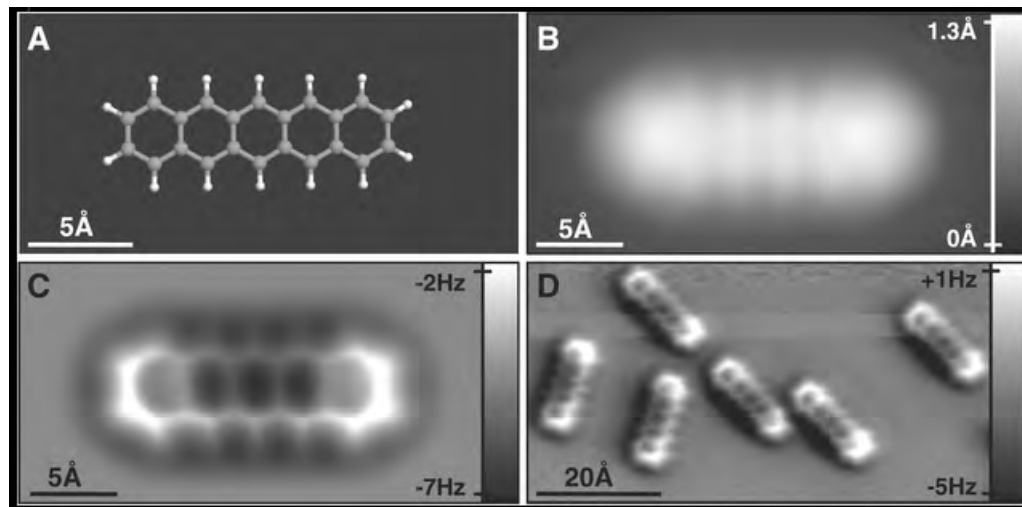


Figure 14. Imágenes de STM y AFM de pentaceno en Cu(111). a) Modelo de la molécula de pentaceno; b) Imagen de STM con corriente constante; c-d) Imágenes de AFM con altura constante adquiridas con una punta modificada con CO [28].

$$F = k_c \delta \quad (17)$$

Donde delta es la deflexión del voladizo (nm) y k_c es la constante de resorte (N/m), que va normalmente de 0.03 to 5 N/m, resultando en fuerzas que van de pN a μ N en condiciones ambientales.

El modo de no contacto (NC-AFM, por sus siglas en inglés) actúa en el régimen de distancias en las que los electrones de átomos vecinos son redistribuidos, induciendo polarización instantánea de las moléculas interactuantes y generando fuerzas de dispersión. Esta interacción dipolo-dipolo induce una fuerza atractiva, la cual se torna dominante por sobre la fuerza de repulsión. En este caso, las fuerzas de van der Waals modifican la amplitud y la fase de la vibración resonante del voladizo induciendo una frecuencia de resonancia efectiva. La fuerza de gradiente experimentada por el voladizo, es decir la derivada de la curva de potencial-distancia en la figura 13, que depende de la distancia punta-muestra, dicta la variación de la constante de resorte efectiva (k_{eff}) del voladizo, la cual a su vez está relacionada con su frecuencia de resonancia efectiva (f_{eff}) como la raíz cuadrada (resolviendo la ecuación del oscilador armónico simple):

$$f_{eff} = (k_{eff}/m)^{1/2} \quad (18)$$

Donde m es la masa del voladizo. Por lo tanto, cambios en la frecuencia de resonancia indican cambios en la distancia punta-muestra. Una fuerza atractiva actuando sobre la punta disminuye la constante de resorte efectiva conforme la punta se acerca a la superficie, resultando en una disminución de la frecuencia de resonancia efectiva [29–33].

2. Resonancia Magnética Nuclear

En la presencia de un campo magnético externo B_z , con intensidad B_0 , un espín nuclear experimenta una interacción cuya energía puede ser representada por el Hamiltoniano:

$$H_1 = -\gamma B_0 I_z \quad (19)$$

Donde I_z es la componente z del operador de momento angular de espín nuclear y γ es la razón giromagnética, una propiedad del núcleo. Siendo la ecuación para el eigenvalor I_z con eigenfunción de onda Ψ :

$$I_z \Psi_m = m \hbar \Psi_m \quad (20)$$

La ecuación de eigenvalor correspondiente al hamiltoniano que nos da los niveles de energía de un espín en un campo magnético es entonces:

$$H_1 \Psi_m = -m \hbar \gamma B_0 \Psi_m \quad (21)$$

Donde m toma los valores entre $-I$ e I en pasos enteros. Para fermiones, $I = \frac{1}{2}$, por lo tanto $m = \pm \frac{1}{2}$, y como consecuencia existen dos posibles energías para el sistema:

$$E_\alpha = -\hbar \gamma B_0, E_\beta = \hbar \gamma B_0 \quad (22)$$

De esta manera, sabiendo que $E = h\nu$, la frecuencia ν del fotón con energía equivalente a la diferencia de energías entre los estados alfa y beta $\Delta E_{\alpha\beta}$ es

$$\nu_{\alpha\beta} = \gamma B_0 / 2\pi \quad (23)$$

Que se define como la frecuencia de Larmor ν_0 , la cual puede también interpretarse como la frecuencia de precesión del vector de magnetización nuclear \mathbf{m} alrededor del eje del campo magnético externo (Fig. 15b). La frecuencia de Larmor está en el rango de radiofrecuencias del espectro electromagnético, y como consecuencia, al irradiar con pulsos de radiofrecuencia en resonancia magnética con los espines nucleares, es posible manipular los cambios de estado.

La magnetización fuera de equilibrio creada por el pulso de radiofrecuencia cercano a la frecuencia de Larmor de los espines nucleares genera el decaimiento libre de inducción (FID, por sus siglas en inglés) que es la señal observable de la resonancia magnética nuclear (NMR, por sus siglas en inglés)(Fig. 15c). El FID generado de la precesión magnética se sujeta a una transformada de Fourier del dominio de tiempo al dominio de frecuencia para obtener el desplazamiento químico determinado por el ambiente químico de cada uno de los núcleos debido al efecto de escudo que variaciones en la distribución electrónica causan sobre la influencia del campo magnético externo en los núcleos.

Los desplazamientos químicos son las unidades que constituyen los espectros de resonancia magnética nuclear e indican las relaciones espaciales entre cada uno de los átomos que conforman el sistema bajo análisis. Junto con los desplazamientos químicos, se usan también restricciones de distancia y ángulos dictadas por i) el efecto nuclear Overhauser, relacionado con la transferencia de magnetización a través del espacio; ii) ángulos diedros, definidos por la planaridad de enlaces y rotaciones de grupos; iii) acoplamiento dipolar residual, interacciones magnéticas entre dos espines resultado de un alineamiento parcial que causa un promediado incompleto de acoplamientos dipolares anisotrópicos; iv) puentes de hidrógeno [34–37].

3. *Dispersión de Rayos X a Ángulos Pequeños*

A ángulos pequeños (0.1- 5.0), la dispersión de rayos X en solución es principalmente elástica y genera información estructural de la muestra basada en los efectos de difracción producidos por

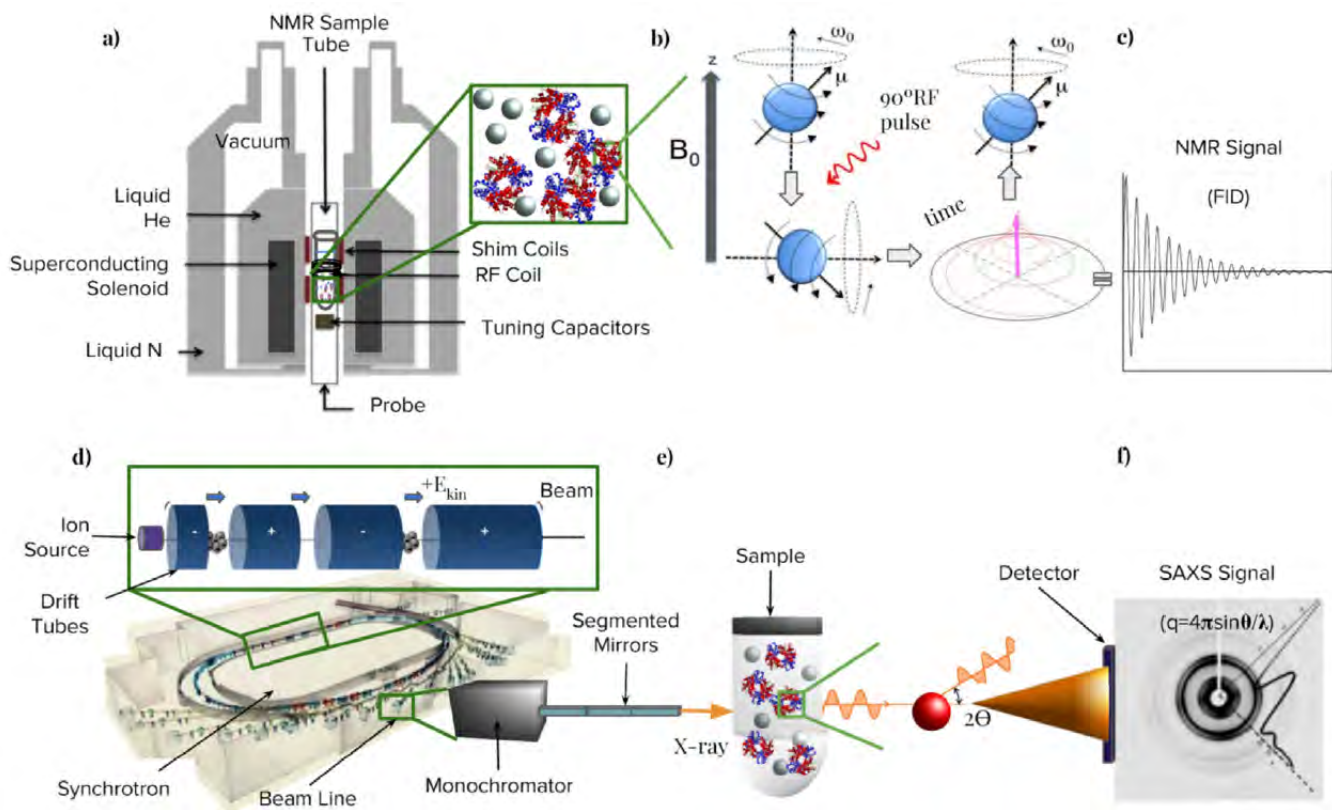


Figure 15. Descripción comparativa de las técnicas NMR y SAXS. a) Espectrómetro de NMR; b) Principio físico de NMR: al irradiar con pulsos de radiofrecuencia en resonancia magnética con los espines nucleares es posible manipular cambios de estado. c) Generación de datos de NMR: la observable de NMR, decaimiento libre de inducción (FID), resulta de la precesión del vector de magnetización del espín nuclear alrededor del eje del campo magnético en un estado fuera de equilibrio; d) Síncrotrón para SAXS conformado por tres etapas principales: producción de partículas, pre-aceleración por un acelerador lineal (linac) y almacenamiento en el anillo del sincrotrón; e) Principio físico de SAXS: la intensidad de rayos X dispersados es función de la transferencia de momento q de cada partícula, que a su vez es función del ángulo de dispersión 2θ y la longitud de onda de radiación incidente; f) Generación de datos de SAXS: la intensidad de dispersión es proporcional a la dispersión de una sola partícula promediada en todas las orientaciones, resultando en una distribución isotrópica con simetría radial.

la interferencia de la dispersión coherente. La diferencia en densidades electrónicas entre el soluto

y el solvente (densidad de exceso de longitud de dispersión) resulta en un patrón de intensidad de dispersión que se detecta y graba. La intensidad de dispersión es función de la transferencia de momento \mathbf{q} de cada partícula, que a su vez es función del ángulo de dispersión 2Θ y de la longitud de onda de la radiación incidente 2λ :

$$q = 4\pi \sin\Theta / \lambda \quad (24)$$

En una muestra monodispersa la intensidad de dispersión es proporcional a la dispersión de una sola partícula promediada en todas las posibles orientaciones, resultando en una distribución isotrópica (con simetría radial). Una transformada de Fourier se aplica a la función de distribución de intensidad (Fig. 2e) para obtener la función de distribución de distancias interatómicas de la partícula $P(r)$, donde r se refiere al vector de distancia interatómica dentro de una molécula. De esta manera, la función de distribución de intensidad $I(q)$ a ángulo cero ($q=0$), que es proporcional a la masa y volumen de la especie muestra, está dada por:

$$I(0) = 4\pi \int_0^{D_{max}} P(r) dr \quad (25)$$

Donde D_{max} es la distancia máxima en la distribución de dispersión, reflejando el tamaño y la simetría interna de la especie en cuestión. La gráfica de Guinier, que representa la función de distribución de intensidad de $I(0)$ y el radio de giro R_g , describe la distribución de masa y brinda información sobre la forma de la especie.

Existe un número creciente de metodologías que conjugan NMR y SAXS para la determinación de estructura de sistemas biológicos en la nanoescala, así como para elucidar sus propiedades dinámicas y dependientes del tiempo. El modelado híbrido NMR-SAXS puede lograrse siguiendo diferentes rutas, entre las que destacan los cálculos ab initio de co-refinamiento. Considerando que en los datos provenientes de SAXS la curva de distancia interatómica es la transformada de Fourier de la función de distribución de intensidad del ángulo de dispersión, éstos contienen la distribución de la densidad atómica, datos que compensan la falta de información traslacional de NMR. En un proceso iterativo usando dinámica molecular, la curva de intensidades de dispersión se simula y se refina con factores

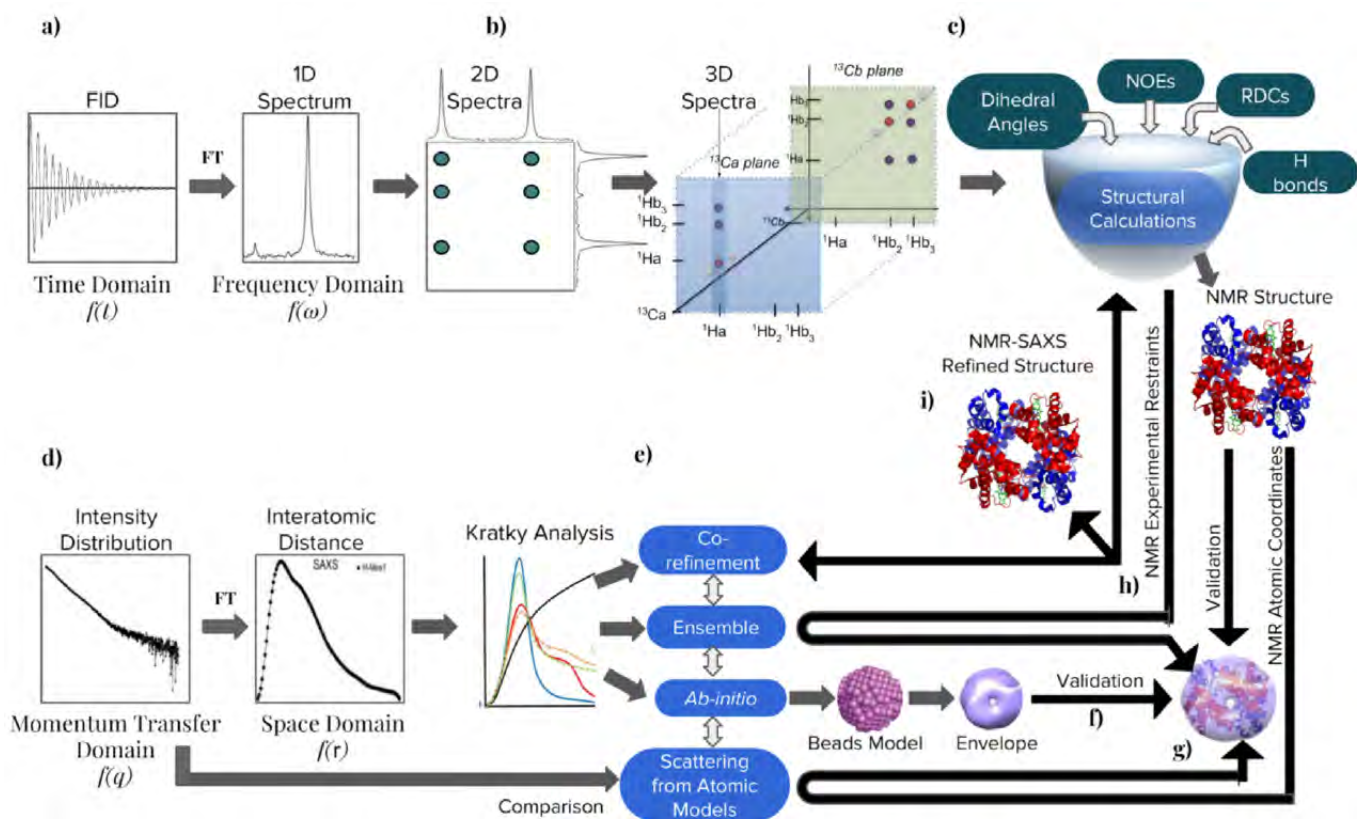


Figure 16.

correctivos que se re-estiman en un ciclo sucesivo hasta que se alcanza la convergencia (Fig. 16) [36, 38–41].

-
- [1] R. L. Johnston, Dalton Trans. , 4193 (2003).
 - [2] H. A. Hussein, J. B. A. Davis, and R. L. Johnston, Phys. Chem. Chem. Phys. **18**, 26133 (2016).
 - [3] P. Gruene, D. M. Rayner, B. Redlich, A. F. G. van der Meer, J. T. Lyon, G. Meijer, and A. Fielicke, Science **321**, 674 (2008), <https://science.sciencemag.org/content/321/5889/674.full.pdf>.
 - [4] L. O. Paz-Borbón, G. Barcaro, A. Fortunelli, and S. V. Levchenko, Phys. Rev. B **85**, 155409 (2012).
 - [5] M. Walter, J. Akola, O. Lopez-Acevedo, P. D. Jadzinsky, G. Calero, C. J. Ackerson, R. L. Whetten, H. Grönbeck, and H. Häkkinen, Proceedings of the National Academy of Sciences **105**, 9157 (2008),

DEL PROTÓN A LOS QUARKS

Dr. Genaro Toledo Sánchez.
Instituto de Física, UNAM.

Introducción

Del protón a los quarks, es un recorrido por algunos de los aspectos más relevantes en el entendimiento de la estructura de la materia, desde lo que consideramos los elementos fundamentales, tales como las partículas y las interacciones, hasta la exploración misma del interior del protón, vía la dispersión inelástica profunda de este con un electrón, y los ingredientes que la cromodinámica cuántica aporta para el entendimiento del mismo. Estas notas surgen del minicurso impartido en la Escuela de Verano en Física 2021, la cual se realizó de manera virtual, por lo que en la dinámica del mismo utilicé el chat de la sesión para que los asistentes participaran con preguntas y/o comentarios. Tratando de emular esto, en cada sección, se formulan preguntas con el propósito de que el lector tenga su propia retroalimentación al concluir el capítulo. Enlaces a sitios que pueden ser de utilidad para extender la información sobre los fenómenos discutidos son presentados al final. La estructura general de este capítulo es la siguiente: En la sección 1, se presentan los elementos básicos que integran al modelo estándar, las partículas elementales y sus interacciones. En la sección 2, abordamos las generalidades de la invariancia de norma local, tomando como ejemplo a los fermiones, para lo cual presentamos la ecuación de Dirac e introducimos el formalismo lagrangeano y vemos las implicaciones de la simetría de norma local $U(1)$. En la Sección 3, ilustramos cómo este concepto, es el ingrediente principal para la formulación del modelo estándar y mencionamos algunos aspectos que aún tienen que ser entendidos. La Sección 4 la dedicamos a estudiar la cromodinámica cuántica y sus principales características, desde la interacción entre gluones, que llevan a tener una imagen más completa de lo que es un hadrón, hasta la evolución de la constante de acoplamiento y sus distintos regímenes, deconfinamiento y libertad asintótica. Concluyendo con una descripción de la dispersión inelástica profunda como medio para explorar el interior de los hadrones. En la sección 5 concluimos presentando las perspectivas enfocada en propiedades hadrónicas y su papel actual en diversos escenarios, los estados multiquarks y exploraciones tanto terrestres como astrofísicas.

1 Las partículas y sus interacciones.

¿Qué pregunta te has hecho respecto a la estructura de la materia? Aún es tiempo de hacerla!

Podemos empezar por preguntarnos ¿Qué es lo que identifica a una partícula? ¿Cuáles de las propiedades físicas que conocemos como la masa, el momento angular total, el espín, la carga, las interacciones que percibe, sus propiedades bajo algún tipo de transformación, etc. son las inherentes a ella? Con ese objetivo, la física de los constituyentes fundamentales de la materia ha pasado por etapas que a la fecha han producido un esquema simple de entendimiento a partir de un mínimo de elementos. Existen 2 grupos principales de elementos: Los *fermiones*, de espín fraccionario y los *bosones*, de espín entero. A nivel fundamental los primeros son llamados quarks y leptones, los cuales son de varios tipos llamados *sabores*. Los segundos son típicamente los mediadores de la interacción entre los primeros, veremos más adelante cómo llegamos a esa afirmación. Los quarks y leptones se muestran en las Figuras 1 y 2 respectivamente, así como sus cargas eléctricas. Los leptones tienen un número cuántico leptónico ($L = 1$) que los caracteriza, mientras que los quarks tienen el número bariónico ($B = 1/3$). También existen las correspondientes antipartículas, las cuales se denotan generalmente con una barra encima del símbolo correspondiente para la partícula con sus números cuánticos de carga eléctrica, B y L de signo opuesto, entre otros.

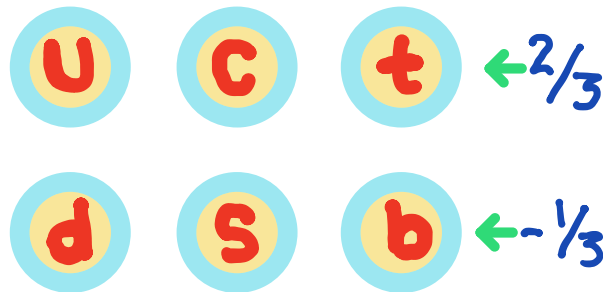


Figure 1: Quarks. Las letras denotan su sabor, y sus cargas eléctricas, en unidades de la carga del positrón se etiquetan para el renglón correspondiente.

Existen más elementos dentro de estos 2 grandes grupos, que se forman a partir de este conjunto mínimo. Grupos formados por un quark y un antiquark se acomodan formando *mesones* y grupos de 3 son llamados *bariones*. Por ejemplo, el protón (p) está formado por los quarks uud y el pión (π^+) por $u\bar{d}$. Es posible que existan estados con más quarks, como los tetraquarks, pentaquarks, etc. En

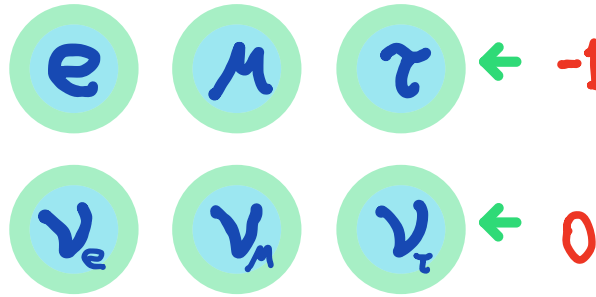


Figure 2: Leptones. Los símbolos denotan su sabor, y sus cargas eléctricas, en unidades de la carga del positrón, se etiquetan para el renglón correspondiente.

general a los grupos de quarks se les denomina *hadrones* .

Un resultado interesante que se puede inferir de los números cuánticos es, por ejemplo, que el protón, siendo el barión más ligero, no se observa que decaiga, lo que se puede interpretar como una manifestación de la conservación del número bariónico. En el sitio <https://pdg.lbl.gov> se pueden encontrar los valores más actualizados de las propiedades de las partículas.

Otra observación interesante es que las masas del protón y el neutrón son muy similares ($m_{proton} = 938.27$ MeV y $m_{neutron} = 939.57$ MeV). Esto sugirió la existencia de la simetría de *Isoespín*, en analogía al espín, el cual tiene asociado un número cuántico correspondiente (I) para los estados con esa propiedad. En este caso, el estado está formado por el protón y el neutrón, correspondiéndoles $I = 1/2$, y distinguimos a uno del otro por su carga. En términos de I , se identifican por su diferente proyección en la dirección 3 (I_3). Esta simetría también se observa a nivel de quarks, en donde podemos identificar la cercanía en masa de los quarks u y d y un poco más grande la del quark s . Esto permitió clasificar a los estados hadrónicos ligeros usando esta propiedad y un número cuántico adicional, la hipercarga fuerte $Y = B + S$, donde S es el número cuántico asociado a los quarks s ($S = -1$). En la Figura 3 se representan los mesones ligeros (compuestos de quarks u , d y s solamente) de espín 1 (vectoriales), en el plano I_3 vs. Y . Una extensión de esta idea, correspondería a incluir a los quarks c , lo que en términos gráficos significa agregar un eje mas, para dar cuenta del contenido de quarks c , convirtiéndose en una representación tridimensional,(ver PDG, Quark model). Sirva esto para mostrar la proliferación de estados hadrónicos. Por otro lado, en la parte leptónica no se observa esta misma característica, dada la naturaleza de sus interacciones, como describimos a continuación.

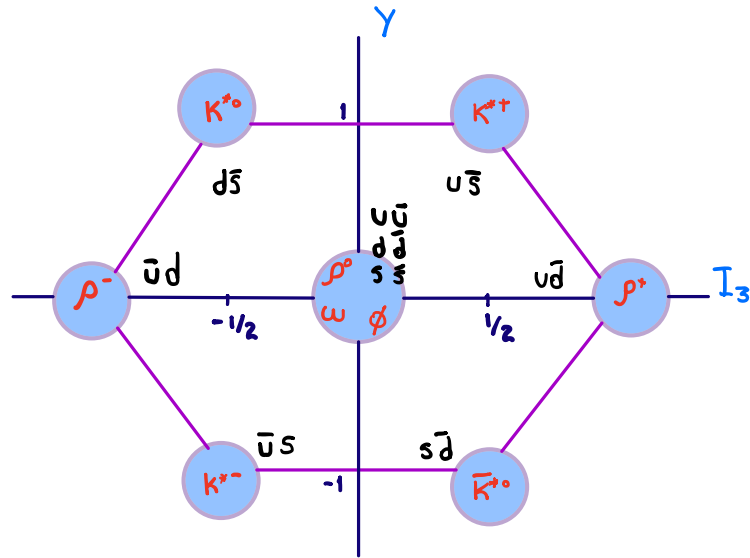


Figure 3: Mesones vectoriales (espín=1) ligeros.

Podemos identificar 4 interacciones fundamentales: La gravitacional, la electromagnética, la fuerte y la débil. Un concepto importante es el de Universalidad. Se considera que las interacciones actúan de igual manera en los distintos sabores de quarks y lo mismo entre leptones. Cada columna de las Figuras 1 y 2, se denominan familias. El modelo estándar considera la existencia de 3 familias.

Antes de continuar, introduzcamos a los llamados *Diagramas de Feynman*, los cuales son una forma práctica de visualizar un proceso de interacción entre partículas. La idea es definir una dirección temporal y, perpendicular a ella, otra espacial. En este plano podemos dibujar líneas que corresponden a la propagación de partículas en una dirección determinada. Usamos líneas sólidas para quarks y leptones y líneas onduladas para los bosones que median las interacciones, como se muestra en la Figura 4. El punto en donde se unen varias líneas se llama **vértice** y una línea intermedia se denomina **propagador**. Utilizando reglas sencillas se puede obtener la información de la probabilidad con que los procesos que involucran estos elementos ocurren.

Interacción electromagnética.

Las partículas cargadas eléctricamente, tales como todos los quarks y los leptones e , μ y τ interactúan a través del intercambio de un fotón (γ). Por ejemplo, la aniquilación de un electrón y un positrón, y la subsecuente creación de un par muón- antimuón, puede visualizarse como debida al intercambio de un fotón,

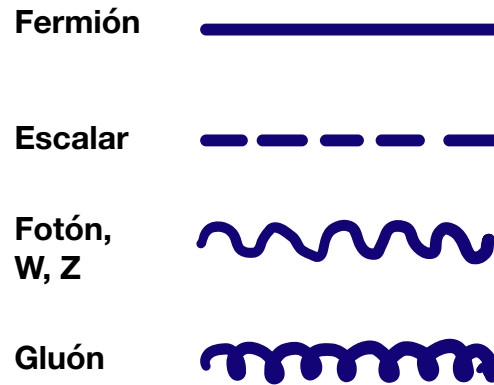


Figure 4: Representación de las distintas partículas en diagramas de Feynman

como lo ilustra la Figura 5.

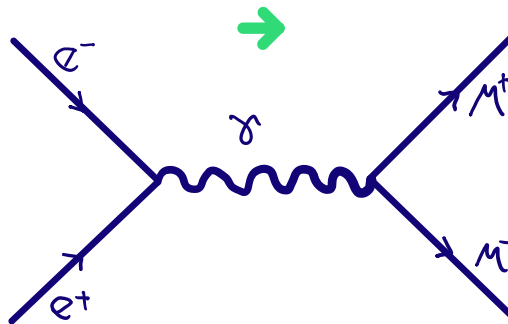


Figure 5: Diagrama de Feynman para la aniquilación de un electrón y un positrón, y la subsecuente creación de un par muón- antimuón, vía el intercambio de un fotón.

El fotón se acopla a un electrón proporcional a su carga, así, para 2 partículas con carga e la constante de estructura fina α caracteriza la magnitud de estas interacciones.

$$\alpha = \frac{e^2}{4\pi\hbar c} = \frac{1}{137} \quad (1)$$

El fotón a nivel fundamental no interactúa directamente consigo mismo, pues tiene carga eléctrica cero. Sin embargo, esto puede ocurrir a través de fluctuaciones de vacío, en las cuales el fotón crea un par de partículas cargadas. La teoría que la

describe es la Electrodinámica cuántica.

Interacción Fuerte.

Los quarks son los únicos fermiones que perciben las interacciones fuertes, los cuales tienen una carga llamada de color, que pueden ser rojo, verde o azul (los leptones no tienen carga de color). Existen 8 partículas que median la interacción, los gluones, a diferencia de la interacción electromagnética en la que solo existe el fotón. Estos pueden cambiar las propiedades de color del quark pero no su sabor, mas aún, los gluones tienen carga de color, por lo que pueden interactuar entre sí. En la Figura 6 se muestra un ejemplo, en donde los quarks, tienen la etiqueta de color. La magnitud correspondiente se describe, en analogía al electromagnetismo, como

$$\alpha_s = \frac{g_s^2}{4\pi} \approx 1 \quad (2)$$

donde g_s es el acoplamiento del gluón con cada quark. Una forma de obtener α_s es a través de la medición de la probabilidad de que una partícula decaiga vía interacción fuerte comparada con el caso en que lo haga electromagnéticamente. α_s es aproximadamente 100 veces mayor que α . Mas adelante veremos en detalle las propiedades de la teoría que la describe, la Cromodinámica cuántica.

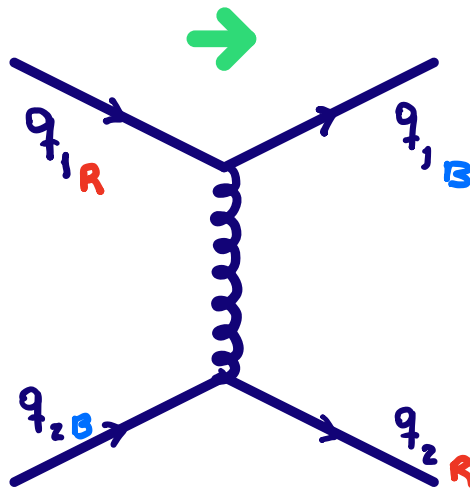


Figure 6: Diagrama de Feynman para la interacción fuerte entre 2 quarks de color rojo y azul a través de un gluón.

Interacción Débil.

Las interacciones débiles afectan tanto a quarks como a leptones. Los mediadores son masivos, dos cargados eléctricamente (W^\pm) y uno neutro (Z^0). La magnitud de la interacción está dada generalmente por la constante de Fermi, la cual es la razón entre la constante de carga y la masa del W

$$G_F \equiv \frac{\sqrt{2}g^2}{8M_W^2} \approx 10^{-5} GeV^{-2} \quad (3)$$

La medición de G_F permitió predecir la masa del W la cual coincidió con las mediciones directas posteriores. El hecho de que G_F sea proporcional a la masa del W y este a su vez muy masivo, es la razón de que se le considere una interacción débil, sin embargo, si realizamos una comparación en analogía a las de los casos electromagnético y fuerte, la constante correspondiente es:

$$\alpha_w = \frac{g^2}{4\pi} \quad (4)$$

De la relación de esta con G_F , vemos que no es tan pequeña. Un caso típico de interacción débil es el decaimiento Beta del neutrón

$$n \rightarrow pe^- \bar{\nu}_e \quad (5)$$

A diferencia de las interacciones fuertes, *las interacciones débiles pueden cambiar el sabor de las partículas*, así, por ejemplo, un quark u del protón es cambiado en un d con acoplamiento g y el neutrino del electrón a un electrón, con la misma magnitud. En el sector de quarks, los eigenestados de masa son mezclas de los eigenestados de interacción. La matriz que parametriza esto es la llamada Matriz de CKM (Cabibbo, Kobayashi, Maskawa), la cual introduce correcciones a las intensidades de las interacciones.

Interacción Gravitacional.

La fuerza debida a la gravedad entre dos masas iguales M está dada por la ley de Newton. No tenemos una versión cuántica para la gravedad, aunque existen muchos intentos, por lo que una forma de comparar su magnitud con las otras interacciones, sin ser del todo correcta, es usar argumentos dimensionales. Así, la escala de Planck es aquella a la cual la fuerza gravitacional se aproxima a uno:

$$M_{Planck} = \sqrt{4\pi\hbar c/G} = 2 \times 10^{19} GeV. \quad (6)$$

Las masas de las partículas elementales son muy pequeñas comparadas con estas magnitudes por lo que las consecuencias de la interacción gravitacional son despreciables en este caso.

¿Cuál de las interacciones te parece más interesante?

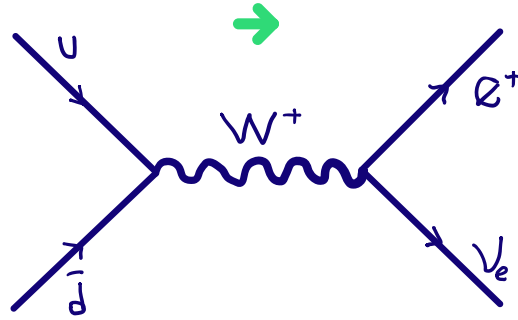


Figure 7: Diagrama de Feynman para un proceso débil en donde un quark u y un $antid$ producen un neutrino y un positrón debido a un W .

2 Ecuación de Dirac e invariancia de norma

¿Las simetrías nos enseñan algo sobre las partículas?

Paul Dirac, propuso una ecuación para estudiar al electrón, la cual es cuántica, incorporando los operadores cuánticos de energía y momento y sus correspondientes reglas de conmutación, y es relativista, cumpliendo la relación entre energía, momento y masa relativista. Usando que dicha ecuación debería satisfacer conservación de la probabilidad esta debería ser lineal en las derivadas temporales. La invariancia ante transformaciones de Lorentz requiere también linealidad en las derivadas espaciales. Esto llevó a Dirac a escribir su ecuación para una partícula fermiónica de masa m y cuádrimomento P^μ como

$$(\gamma \cdot P - m)\psi = 0, \quad (7)$$

donde ψ es una función de 4 componentes y γ son las matrices de Dirac $\gamma^\mu = (\gamma^0, \gamma^i)$, $i=1,2,3$.

$$\gamma^0 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \quad \gamma^i = \begin{pmatrix} 0 & \sigma_i \\ -\sigma_i & 0 \end{pmatrix} \quad (8)$$

con σ_i las matrices de Pauli.

La notación relativista que utilizaremos para la descripción matemática es la siguiente: Un cuádrivector es de la forma $a^\mu = (a^0, a^1, a^2, a^3) = (a^0, \mathbf{a})$, el tensor métrico por convención es $g_{\mu\nu} = \text{diag}(1, -1, -1, -1)$ y el producto de 2 cuádrivectores es $a^\mu b_\mu = a^0 b_0 - a^1 b_1 - a^2 b_2 - a^3 b_3$.

Las simetrías de la naturaleza dictan en gran medida la forma de describirla. Así, resultados como la conservación de la energía y la invariancia ante transformaciones de Lorentz forman parte del esquema básico de cualquier teoría que pretenda explicar un fenómeno físico. La formulación usando Lagrangeanos ofrece la ventaja de encontrar de manera fácil las corrientes conservadas de cualquier teoría a la vez que muestra en su construcción la invariancia de Lorentz.

El Lagrangeano (Usaremos este término para referirnos al Lagrangeano (L) o a la densidad lagrangeana indistintamente \mathcal{L}) está dado por

$$L = \int d^3x \mathcal{L}(\phi, \partial_\mu \phi) \quad (9)$$

Al minimizar la acción S ($S = \int_{t_1}^{t_2} L dt$) obtenemos las ecuaciones de Euler-Lagrange

$$\frac{\partial \mathcal{L}}{\partial \phi} - \partial_\mu \frac{\partial \mathcal{L}}{\partial (\partial^\mu \phi)} = 0 \quad (10)$$

donde ϕ es una variable canónica. La ecuación de movimiento correspondiente se obtiene al poner explícitamente a \mathcal{L} . En la formulación cuántica, las observables son calculadas usando el factor de probabilidad de ocurrencia de un evento, que corresponde a la función de onda al cuadrado $|\phi|^2$. Esto deja como opción que la función de onda esté determinada hasta una fase, $\phi \rightarrow e^{-i\alpha} \phi$. El que la teoría sea invariante ante esta transformación es llamada **invariancia de norma**. Si el parámetro α es constante corresponde a una *invariancia global* y si $\alpha = \alpha(x)$, es decir depende del espacio-tiempo, se denomina *invariancia de norma local*.

En física de partículas elementales, el Lagrangeano se puede escribir en función de los campos asociados a las partículas elementales de la teoría y, como en electrodinámica clásica, este Lagrangeano contendrá toda la física correspondiente. Es decir, escribir el Lagrangeano de la teoría es equivalente a formularla.

Consideremos ahora el Lagrangeano de una partícula *fermiónica libre* ψ de masa m

$$\mathcal{L}_0 = \bar{\psi}(x)(i\gamma^\mu \partial_\mu - m)\psi(x) \quad (11)$$

donde γ^μ son las matrices de Dirac mencionadas anteriormente. Veamos qué pasa cuando exigimos invariancia bajo una transformación de norma local.

$$\psi(x) \rightarrow \psi'(x) = e^{-i\alpha(x)}\psi \quad \bar{\psi}(x) \rightarrow \bar{\psi}'(x) = e^{i\alpha(x)}\bar{\psi}. \quad (12)$$

Este no es invariante a menos que se introduzca un nuevo campo, A^μ , llamado *campo de norma* y la derivada parcial se sustituya por la llamada derivada covariante $\partial_\mu \rightarrow \partial_\mu + ieA_\mu$, con lo cual el Lagrangeano se reescribe como

$$\mathcal{L} = \bar{\psi}(x)i\gamma^\mu(\partial_\mu + ieA_\mu)\psi(x) - m\bar{\psi}(x)\psi(x) - \frac{1}{4}F^{\mu\nu}F_{\mu\nu} \quad (13)$$

donde A_μ se transforma como

$$A^\mu \rightarrow A'^\mu = A^\mu + \partial^\mu \alpha(x)/e \quad (14)$$

y

$$F^{\mu\nu} = \partial^\mu A^\nu - \partial^\nu A^\mu. \quad (15)$$

La primera parte corresponde al término cinético y el segundo al de potencial o interacción. Dado este Lagrangeano las ecuaciones de movimiento corresponden a las ecuaciones de Maxwell. Este resultado tiene como consecuencia principal, la aparición de un campo que esencialmente tiene la función de transmitir la interacción. En la Tabla 2, se muestran los términos y la información física que les corresponde.

$\psi(x)i\gamma^\mu\partial_\mu\psi(x)$	Término cinético de ψ
$m\psi(x)\psi(x)$	Término de masa de ψ
$\psi(x)e\gamma^\mu A_\mu\psi(x)$	Término de interacción entre $\psi(x)$, $\psi(x)$ y A_μ
$F^{\mu\nu}F_{\mu\nu}$	Término cinético de A_μ

Tabla 2. Descripción de los términos del Lagrangeano (Eqn. 13)

El último término se incluye también por satisfacer las condiciones de invariancia de norma y de Lorentz. Note que no existe un término de masa para A_μ , que sería de la forma $m_A A^\mu A_\mu$ pues rompería la simetría de norma, de donde se tiene que $m_A = 0$. En un caso más generalizado se puede mostrar que el fenómeno de interacción débil y fuerte son un resultado de la invariancia de norma local, con sus correspondientes grupos de simetría. En general para grupos de simetría $SU(n)$ hay $n^2 - 1$ bosones de norma, asociados. Los correspondientes campos de norma en la parte débil con grupo $SU(2)$ son los W^\pm , Z^0 y para la interacción fuerte, con grupo de norma $SU(3)$, 8 gluones. Similarmente se espera que la gravedad tenga asociado el gravitón aunque no se ha demostrado.

¿De dónde surgen los mediadores de la interacción?

3 El Modelo Estándar

¿Cómo calculamos la probabilidad de que ocurra un proceso?

A finales de 1960 Weinberg, Glashow y Salam formularon de qué manera se podía tratar a la interacción electromagnética y la débil como aspectos diferentes de una teoría electrodébil. Esta sería evidente a energías en las que el cambio en el

momento transferido al cuadrado (q^2) fuera mayor que 10^4 GeV. A bajas energías esta simetría está rota. De los 4 bosones de norma involucrados 3 son masivos (W^\pm, Z^0) y uno es no masivo (fotón). Así, al comparar electromagnetismo con la interacción débil esta última es de corto alcance y muy pequeña. A continuación discutimos cómo se constituye el modelo electrodébil, para eso recurriremos a las ideas de invariancia de norma mencionadas anteriormente.

- Todas las partículas parecen tener una invariancia de fase que llamaremos U(1) la cual podemos indirectamente asociar con electromagnetismo, esto requiere que exista un campo de norma sin masa al que denotaremos B^μ .
- Hay en las partículas una segunda simetría de norma llamada SU(2), que corresponde a rotaciones en un espacio similar al del espín llamado isospín, la cual tiene asociado 3 bosones de norma W_1^μ, W_2^μ y W_3^μ .
- Una tercera simetría interna es llamada SU(3), con 8 bosones de norma asociados llamados gluones. La teoría que describe la interacción vía gluones se llama cromodinámica cuántica (QCD) en donde las partículas llevan una *carga de color*.

El modelo de Weinberg, Glashow y Salam se refiere a la parte de U(1) y SU(2). Utiliza el proceso de rompimiento espontáneo de la simetría para generar las masas de los campos de norma mientras se mantiene la invariancia de norma. Esto se realiza con la ayuda de un doblete de partículas escalares complejas, llamadas de Higgs. Después del rompimiento, 3 bosones adquieren masa (W^\pm, Z^0) y otro permanece sin masa (A_μ). Estos son combinaciones de los primeros de la forma

$$W^\pm = \frac{1}{\sqrt{2}}(W_\mu^1 \pm W_\mu^2) \quad (16)$$

$$Z_\mu = W_\mu^3 \cos\theta - B_\mu \sin\theta \quad (17)$$

$$A_\mu = W_\mu^3 \sin\theta + B_\mu \cos\theta \quad (18)$$

donde θ es el ángulo de mezcla débil (o ángulo de Weinberg) el cual experimentalmente se encuentra ser $\sin^2\theta \approx 0.223$. Las masas de los quarks y leptones también se generan por el mismo mecanismo. Lo que hoy llamamos modelo estándar está constituido por el producto de las 3 simetrías internas $(U(1) \times SU(2))_{\text{electrodébil}} \times SU(3)$ donde la parte de color es una simetría exacta.

Toda teoría física tiene sus raíces en observaciones experimentales y su validéz está basada en la correspondencia con estas últimas. Además de esto, la teoría debe tener un poder predictivo, tal es el caso del modelo electrodébil que predijo

la existencia de los bosones de norma masivos, el no haberlos encontrado habría significado modificarla. Así, la física de partículas elementales es dirigida por la información obtenida en los experimentos, y para eso debemos entender su funcionamiento.

Aunque el modelo estándar de las partículas elementales ha mostrado ser exitosa en describir la fenomenología hasta ahora observada, existen muchas preguntas para las cuales aún no tenemos una respuesta y que son objeto de estudio. A continuación listamos algunas de ellas:

- ¿Existe una interacción universal a altas energías de donde se desprenden las que conocemos?
- ¿Por qué las masas de las partículas tienen esos valores?
- ¿Hay más partículas?
- ¿Por qué el ángulo de mezcla débil tiene ese valor?
- ¿Existe una simetría entre fermiones y bosones?
- ¿Por qué vivimos en un mundo de 4 dimensiones?

Hay otro tipo de preguntas, si bien no de carácter fundamental, si indispensables para entender mejor y desarrollar nuevas estrategias para seguir explorando la fenomenología de las partículas:

- ¿Cómo podemos acelerar partículas a mayores energías?
- ¿Cómo incrementamos la precisión de lo ya medido para ver que corresponde con la teoría conocida?
- ¿Cómo se distribuyen los quarks dentro de los hadrones?
- ¿Cómo medimos propiedades de partículas que viven alrededor de 10^{-23} segundos?

4 La interacción fuerte

¿Cómo exploramos a los hadrones?

El estudio de las propiedades del protón y posteriormente el neutrón en procesos de dispersión, marcaron el inicio de la física hadrónica, la observación de

nuevos estados que compartían algunas características con ellos llevaron a la construcción del llamado modelo de quarks, así, estados extraños (con contenido de quarks s), encantados (con contenido de quarks c) y bellos (con contenido de quarks b) se han vuelto parte de la fenomenología estándar actual, de la cual se ha podido aprender no solo sobre la interacción fuerte, si no también sobre las débiles y electromagnéticas.

La interacción fuerte, está regida por la simetría $SU(3)$ la cual no está rota, tiene características, como el fenómeno de confinamiento y el asociado incremento de la intensidad de la interacción a bajas energías, y por otra parte la libertad asintótica a altas energías. En esta sección describimos de manera general a la cromodinámica cuántica y mostramos ejemplos que corroboran esta descripción. La interacción fuerte se manifiesta a nivel fundamental entre quarks, que son partículas de espín $1/2$ y carga eléctrica fraccionaria y tiene un número cuántico o carga de color que puede ser R , G o B (por las siglas en inglés de rojo, verde y azul). Note que estas son solo etiquetas y no tienen que ver con los colores en el sentido de la óptica. La simetría de norma local $SU(3)$ asociada al color, requiere la existencia de 8 campos de norma, los *gluones*, los cuales no tienen masa. El Lagrangeano que describe a los quarks, denotados por $\psi_{q,b}$ donde q y b son los índices de sabor y color respectivamente, y los gluones, denotados por G_μ^C , se puede poner de la siguiente manera:

$$\mathcal{L} = \sum_q \bar{\psi}_{q,a} \left(i\gamma^\mu \partial_\mu \delta_{ab} - g_s \gamma^\mu t_{ab}^C G_\mu^C - m_q \delta_{ab} \right) \psi_{q,b} - \frac{1}{4} F_{\mu\nu}^G F^{G\mu\nu}, \quad (19)$$

donde

$$F_{\mu\nu}^G = \partial_\mu G_\nu^A - \partial_\nu G_\mu^A - g_s f_{ABC} G_\mu^B G_\nu^C, \quad (20)$$

y $A, B, C = 1, 2, \dots, 8$. Los generadores del grupo satisfacen:

$$[t^A, t^B] = i f_{ABC} t^C, \quad (21)$$

y las representaciones de estos corresponden a las matrices λ^i que cumplen con las relaciones (aquí $a, b, c = 1, \dots, 8$):

$$[\lambda^a, \lambda^b] = 2i f^{abc} \lambda^c; \quad Tr[\lambda^a, \lambda^b] = 2\delta_{ab} \quad (22)$$

Con estos elementos básicos, es posible ya ver parte de la fenomenología que la hace diferente a la electrodinámica cuántica. Por ejemplo, en el término cinético de los gluones $F_{\mu\nu}^G F^{G\mu\nu}$, se puede ver que tendremos productos de 2, 3 y 4 campos gluónicos, estos 2 últimos asociados a la parte no-abeliana del grupo de simetría, así, tendremos vértices de la forma $g \rightarrow q + \bar{q}$, $g \rightarrow g + g$ y $g + g \rightarrow g + g$, como se muestran en la Figura 8.

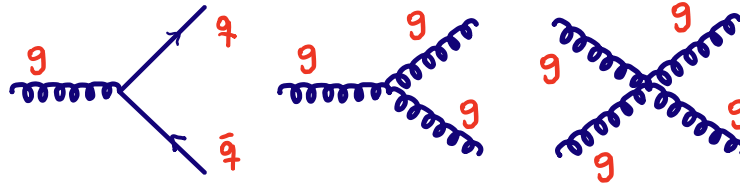


Figure 8: Vértices básicos de la interacción fuerte.

Con estos ingredientes, es natural identificar que la estructura del protón es mas complicada que la que utilizamos con la imagen de 3 quarks ligados por resortes. La estructura completa requiere de considerar todas las posibles formas de ligar a dichos quarks vía los diagramas básicos anteriores. Por ejemplo. Un gluón emitido de un quark puede ir a un vértice de 4 gluones, los que a su vez pueden repetir el mismo mecanismo, antes de conectarse con otro de los quarks. Esta complejidad no es aparente, si consideramos que la suma de estas probabilidades es proporcional a potencias de la constante de acoplamiento g_s , la cual en el régimen de bajas energías no es pequeña, como discutiremos más adelante, y por lo tanto las contribuciones no son necesariamente despreciables. En la Figura 9 mostramos una de estas posibles realizaciones.

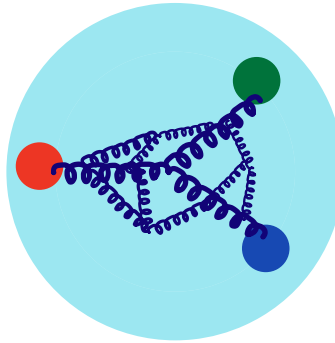


Figure 9: Una posible realización de la conexión entre quarks vía gluones.

El otro aspecto importante es la evolución en energía de la constante de acoplamiento fuerte. En 2004, se otorgó el premio Nobel a David J. Gross, H. David Politzer y Frank Wilczek por el descubrimiento de la libertad asintótica en la teoría de la interacción fuerte. Esta la podemos resumir con la siguiente

ecuación:

$$\alpha_s(Q^2) = \frac{\alpha_s(\mu^2)}{1 + \frac{\alpha_s(\mu^2)}{12\pi}(33 - 2n_f)\log(Q^2/\mu^2)} \quad (23)$$

Note que el denominador siempre crece para el número de sabores n_f menores a $33/2$. El número actual de sabores conocidos es 6. El parámetro μ^2 corresponde a un valor al cual se ha realizado la medición de este acoplamiento. Así, lo que se predice es la evolución como función de la energía al cuadrado Q^2 . En la Figura 10, tomada del Particle Data Group, se muestra el comportamiento teórico esperado de la constante de acoplamiento ($\alpha_s = g_s^2/4\pi$) como función de la energía, y los resultados experimentales, representados por símbolos, los cuales concuerdan muy bien entre sí, correlacionando los extremos con los regímenes de confinamiento y libertad asintótica ya mencionados.

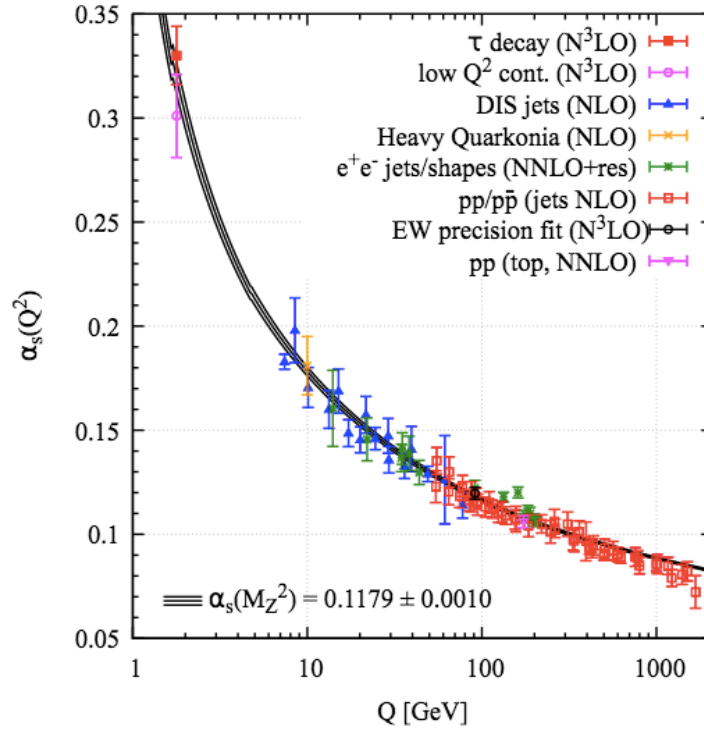


Figure 10: Evolucion en energía de la constante de acoplamiento fuerte

Una de las herramientas para explorar el interior de los hadrones es la llamada dispersión inelástica profunda (DIS por sus siglas en inglés). En ella, una sonda electromagnética explora el interior del hadrón. Para ello, por ejemplo, un electrón se hace colisionar con un protón, intercambiando un fotón, dependiendo de la energía intercambiada, este fotón es capaz de ver la distribución de los quarks y

determinar el radio de esta distribución (también llamada radio de carga, pues solo ve a las partículas cargadas eléctricamente). Otro aspecto importante de este proceso es que, por su naturaleza, no ve a los gluones y por lo tanto el deficit observado del momento de los quarks, para dar cuenta del momento total del protón, dieron indicaciones de la presencia de algo mas que no era detectado por los fotones, estos son los gluones y llevan una fracción cercana al 50% del momento del protón. En la Figura 11, se ilustra la exploración del interior del protón, como una suma de las posibilidades de que el fotón golpee a uno de los quarks en el interior, el cual lleva una fracción x del momento total del protón.

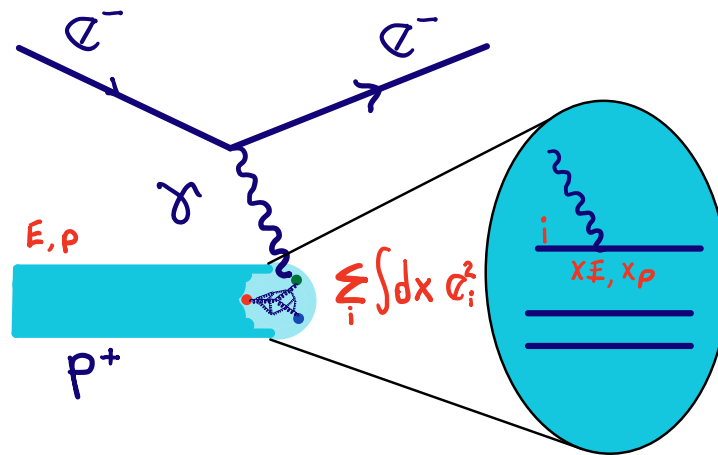


Figure 11: Dispersión inelástica profunda, vista como suma de procesos internos de incidencia del fotón en quarks del interior del hadrón.

La búsqueda de la región de deconfinamiento, es realizado en los grandes aceleradores a través de colisiones a muy altas energías, considerando iones pesados, el estado llamado plasma de quarks y gluones en el que estos prácticamente se comportan como si estuvieran libres ha mostrado ser más rico en fenomenología que lo que se pensó inicialmente, pues fenómenos colectivos, no considerados previamente, parecen ser relevantes. Estos estados de deconfinamiento también podrían estar presentes en las estrellas de neutrones, cuya parte central alcanza densidades enormes comparadas con las densidades nucleares normales. Así, una de las áreas de investigación se centran en entender como ocurre la transición del estado confinado al deconfinado o viceversa, como función de la densidad. En los estudios que hemos realizado sobre el tema, hemos utilizado un modelo que simula numéricamente el comportamiento de los quarks conforme la densidad se incrementa. En este modelo los quarks a bajas densidades están confinados mientras que a altas densidades están esencialmente libres. Algunos resultados que

hemos obtenido son qué bajo condiciones de densidad la aparición de quarks s (extraños) es energéticamente favorable y determinado que la materia compuesta de esa manera es más estable que la materia con quarks u y d solamente. Las implicaciones en las propiedades de la estrella son importantes, pues permitirían la existencia de estrellas con radios menores a los 10 km con una masa de alrededor de 1.4 veces la del Sol.

5 Perspectivas

Los hadrones descubiertos hasta ahora, siguen ofreciendo retos tanto teóricos como experimentales. Ya mencionamos el interés en la distribución interna de los quarks, el cual depende del sabor del quark y se ha explorado muy poco el caso de los hadrones con quarks pesados. En la parte teórica, los estados hadrónicos resonantes son escenarios ideales para probar modelos que intentan capturar las características más relevantes de la interacción fuerte, así como sus propiedades generales. Por ejemplo, los mesones vectoriales (espín 1) tienen tiempos de vida extremadamente cortos, tales que las técnicas experimentales para medir sus propiedades electromagnéticas han sido poco exitosas. La carga eléctrica es determinada por conservación de la misma mientras que el momento dipolar magnético (MDM) es todavía un gran reto. Usualmente, para partículas estables, el método de precesión del espín en un campo magnético constante, dados los estados iniciales y finales de polarización, permite determinar el momento dipolar magnético. Una pregunta es: ¿Como podemos hacerle eso a algo que tan pronto se crea ya ha desaparecido? Una opción alternativa es observar el proceso en el que se emite un fotón ya que este lleva la información electromagnética de donde fue emitido. Así, hemos propuesto varios procesos en los que al observar la distribución angular y energética de estos fotones, se puede ser sensible al valor del MDM. El valor experimental, además de proporcionar un mejor conocimiento de las partículas, al compararla con la predicción teórica del MDM permitirá saber si esta última captura los efectos relevantes, desechando a aquellos modelos que no sean capaces de hacerlo.

Podemos mencionar que la física de mediciones de precisión, también llamada la frontera de la intensidad en física de partículas, es un escenario que, motivado por la búsqueda de física mas allá del modelo estándar, requiere en muchos casos un conocimiento profundo de las llamadas contribuciones hadrónicas, como es el caso del momento magnético anómalo del muón, en el cual los errores asociados a la predicción teórica está fuertemente determinada por las contribuciones hadrónicas.

En la exploración de los estados hadrónicos, empezando con el protón, se ha

recorrido un camino que nos ha llevado a identificar la presencia de los quarks y sus distintas características. Los gluones ahora son parte fundamental de la descripción de la interacción fuerte y se han abierto puertas para estudiar nuevos estados hadrónicos, tanto con contenido de sabores pesados, como con un número mayor a 3 de quarks. Observables en colisionadores y en objetos astrofísicos ponen a prueba nuestro conocimiento de ellos. También los llamados estados exóticos están proliferando en su detección, llamando a un conocimiento más profundo de la interacción fuerte en sistemas colectivos. Esto nos muestra que la búsqueda de nuevas herramientas teóricas y experimentales que nos ayuden a dar respuesta a nuestras interrogantes sigue en marcha.

Bibliografía y sitios de interés.

- D. H. Perkins *Introduction to High Energy Physics*. Addison-Wesley 1987.
- G. Kane *Modern elementary Particle Physics*. Addison-Wesley 1987.
- K. Gottfried, V. F. Weisskopf *Concepts of Particle Physics I*, Oxford University Press 1986.
- F. Halzen and A. D. Martin *Quarks and leptons: An introductory course on modern particle physics*, John Wiley and Sons. 1984

<http://148.247.8.116/> *Division de particulas y campos de la SMF*.

<http://inspirehep.net> *Information de altas energías*.

<http://www.desy.de> *Experimento DESY*.

<http://arxiv.org> *Base de datos*

<http://www.cern.ch> *CERN*.

<http://www.fnal.gov> *Fermilab*.

<http://pdg.lbl.gov> *Particle Data Group*.

<http://www.fisica.unam.mx/grupos/altasenergias> *Altas energías, IFUNAM*.

Láseres: teoría y un experimento

Mayo Villagrán Muniz

Laboratorio de Fotofísica

Centro de Ciencias Aplicadas y Desarrollo Tecnológico

Universidad Nacional Autónoma de México

mayo@aleph.cinstrum.unam.mx

RESUMEN:

Se presenta en forma resumida la teoría básica del funcionamiento de los láseres, se detallan un tipo de láser continuo (He-Ne) y dos pulsados, el de Rubí porque fue el primer láser en operar y el otro más moderno (Nd:YAG) con el que se realizará un experimento en el Laboratorio de Óptica Láser del edificio Tlahuizcalpan (“lugar donde nace la luz”) de la Facultad de Ciencias de la UNAM. La idea de esta parte de la Escuela, es que los estudiantes entiendan los conceptos básicos asociados con los láseres y que los vean en la práctica en un experimento.

INTRODUCCION:

En todos los libros de óptica, cuando se refieren a la fuente de luz utilizada en los experimentos o en el desarrollo de la teoría se cita: “sea una fuente puntual monocromática y coherente”. Solo a partir de la invención del láser en 1960 podemos contar con una fuente de luz que cumple con estas características, permitiendo obtener mejor claridad tanto en los conceptos teóricos como en los experimentos, sean estos de óptica geométrica (rayos) como de óptica física (ondas).

Paralelamente a las innovaciones en los láseres, se han desarrollado numerosas técnicas que los utilizan como herramienta, nuevos materiales para la optimización de ellos mismos, nuevos componentes que integran elementos ópticos y electrónicos (opto-electrónica o electro-óptica) y nuevos elementos mecánicos que fusionan las tres tecnologías (óptica, electrónica y mecánica) en una sola.

En los cursos de óptica y de física moderna a nivel de Licenciatura, generalmente no hay tiempo para presentar estas nuevas tecnologías, ya que se necesita tener una base de mecánica cuántica y electromagnetismo, para en principio, entender el funcionamiento de los láseres dado que “son sistemas cuánticos”.

Se comenzará explicando ¿qué es un láser?, ¿qué contienen?, explicaremos esto de la emisión de luz con fórmulas, los tipos de láseres y los requerimientos teóricos para que puedan operar. Detallaremos uno de los láseres más usuales en los laboratorios de enseñanza (He-Ne), el primer láser (Rubí) y otro (Nd:YAG) que usaremos en un experimento.

¿QUE ES UN LÁSER?

La palabra LASER es un acrónimo del inglés, que significa *Light Amplification by Stimulated Emission of Radiation* que en español sería: Amplificación de Luz por Emisión Estimulada de Radiación.

Lo que es la luz, más o menos ya lo entendemos al igual que lo que significa amplificar. En las últimas tres letras de la palabra Láser (SER), está el corazón por el cual funcionan todos ellos, este nuevo mecanismo de “emisión estimulada de radiación” fue propuesto teóricamente por A. Einstein en 1917, estas ecuaciones para la absorción y emisión de radiación en los átomos, será lo primero que revisaremos.

Un láser es una fuente de luz, pero con características distintas de las fuentes de luz convencionales.

La luz láser es *monocromática*, de un solo color, es decir, la onda electromagnética es de una sola frecuencia, o los fotones tienen una misma longitud de onda.

La luz es *coherente*, las ondas (o los fotones), está en fase.

La luz láser es *direccional*, se propaga en una sola dirección con poca divergencia.

Hay otras características de los láseres, como por ejemplo, la brillantez (debido a la direccionalidad), las altas potencias que pueden alcanzar, pueden ser continuos o pulsados, la duración de los pulsos, el tamaño, etc. Si bien cada una de estas características tienen un sin número de aplicaciones, en lo que sigue trataremos de entender la física del láser y veremos con detalle sólo tres de ellos, dos de los cuales usaremos para una aplicación en el laboratorio.

¿QUÉ CONTIENE?

La mayoría de los láseres, prácticamente contienen tres cosas:

1) Un “medio activo”, es decir “lo que emite luz”, todas las fuentes de luz natural o artificial poseen un medio que emite esta luz. En un foco, el filamento que se calienta por el paso de la corriente, o si es de gases, los gases que emiten. Si el láser es una fuente de luz, necesariamente también tendrá que tener un medio activo.

2) Una fuente de excitación. Esto suena también lógico y común a todas las fuentes de luz, si no excitamos, no va a haber decaimiento de esa energía en forma de luz. Además, como la palabra láser tiene la “A” de amplificación, al igual que cualquier otro amplificador, habrá que entregarle energía.

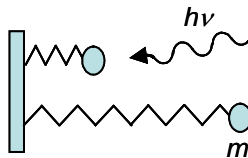
c) Una cavidad resonante. Esta es una diferencia básica que tienen los láseres con las otras fuentes de luz. Esta cavidad simplemente pueden ser un par de espejos colocados en los extremos del medio activo, que reinyectarán parte de la luz que él emite. Las cavidades de los láseres definen la direccionalidad y los modos posibles de oscilación, pero no entraremos en detalle en este punto.

A principios del siglo pasado los físicos tuvieron “la catástrofe del ultravioleta”, donde no se podía explicar la curva de emisión de un cuerpo negro en la región de longitudes de onda corta. Planck la solucionó cuantificando la energía y Einstein, basado en ese postulado, propuso un nuevo mecanismo para la emisión de luz.

COEFICIENTES DE EINSTEIN

Es posible obtener la expresión de radiación de cuerpo considerando la interacción entre los átomos de las paredes de una cavidad y las ondas electromagnéticas. Pensemos en los átomos de las paredes como osciladores que pueden absorber y emitir energía electromagnética, en cantidades que son múltiplos de $h\nu$, es decir, que asumimos la hipótesis cuántica de Planck.

Por ejemplo, para absorción:

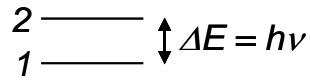


La amplitud del oscilador se ha incrementado.

Einstein identificó tres procesos por medio de los cuales los átomos de la pared pueden interactuar con los campos electromagnéticos.

- a) Emisión Espontánea A_{21}
- b) Absorción B_{12}
- c) Emisión estimulada B_{21}

Asumamos un grupo de osciladores que pueden tener dos posibles estados de energía separados por una energía $h\nu$. Llamemos N_2 a la densidad de osciladores atómicos en el estado 2 y N_1 a la densidad de osciladores en el estado 1



Podemos tener:

- a) EMISIÓN ESPONTÁNEA. Los osciladores atómicos en el estado de alta energía (2) pueden emitir un cuanto de energía electromagnética (fotón) $E = h\nu$ y decaer al estado (1).

La probabilidad de que esto ocurra está dada por un *coeficiente de emisión espontánea* A_{21} . Si la densidad de población del estado 2 es N_2 , la emisión espontánea causará un decaimiento de la densidad de este estado dada por

$$\left. \frac{dN_2}{dt} \right|_{\text{espont}} = -A_{21}N_2 \quad (1)$$

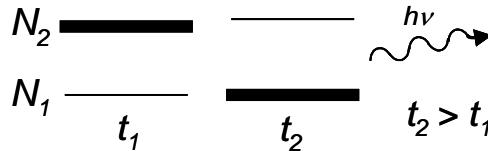


Fig 1

Si ocurre este proceso, la población del estado 2 decaerá como función del tiempo como:

$$\frac{1}{N_2} \frac{dN_2}{dt} = -A_{21} \rightarrow N_2(t) = N_2(t=0) \exp[-A_{21}t] \quad (2)$$

Notemos que la constante de decaimiento de la población del estado 2 en el estado 1 es

$$\Gamma_{21} = \frac{1}{A_{21}} \quad (3)$$

Típicamente los “tiempos de vida” de los estados excitados es de $\Gamma \cong 10^{-8} \text{ s}$.

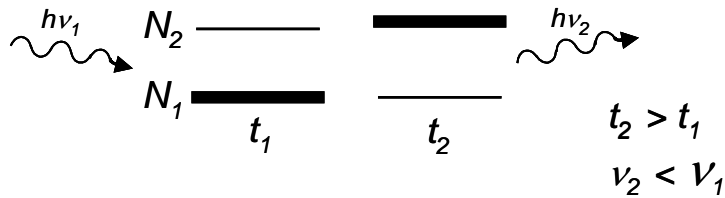
Algunos estados que tienen tiempos de vida muy grande son llamados “estados metaestables”. El estado de energía más bajo es el “estado fundamental” y tiene un tiempo de vida infinito porque no puede decaer a otro estado más bajo.

- b) ABSORCIÓN. Un oscilador en un estado de energía bajo (el 1 en nuestro caso), puede absorber un cuanto de energía (fotón) del campo de la cavidad e incrementar su energía en $h\nu$, finalizando en el estado 2. El ritmo al cual este proceso toma lugar es proporcional al número de átomos en el estado 1 y a la densidad de energía del campo en la cavidad $\rho(\nu)$.

$$\left. \frac{dN_2}{dt} \right|_{\text{ABSOR}} = B_{12}N_1\rho(\nu) = -\frac{dN_1}{dt} \quad (4)$$

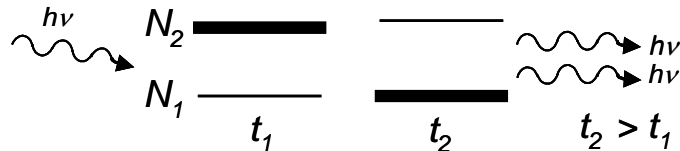
Así como N_2 se incrementa, N_1 decrece

Fig. 2



c) EMISIÓN ESTIMULADA. Es el proceso inverso a la absorción y es el fenómeno básico que permite la amplificación de radiación, es decir que “hace posible el funcionamiento de los láseres”.

Fig. 3



$$\left. \frac{dN_2}{dt} \right|_{EM. ESTI} = -B_{21}N_2\rho(\nu) = -\frac{dN_1}{dt} \quad (5)$$

La emisión estimulada depende de la densidad de átomos en el estado excitado y de la fuerza del campo.

El cuanto de energía emitida tiene la misma frecuencia, fase, polarización y dirección de propagación que el campo incidente.

RELACIÓN ENTRE LOS COEFICIENTES DE EINSTEIN

Einstein usó este proceso para calcular el espectro de energía de un cuerpo negro.

En equilibrio termodinámico

$$-\frac{dN_1}{dt} = \frac{dN_2}{dt} = 0 \quad (6)$$

$$0 = \frac{dN_2}{dt} = -A_{21}N_2 + B_{12}N_1\rho(\nu) - B_{21}N_2\rho(\nu) \quad (7)$$

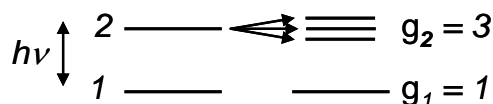
$$\frac{N_2}{N_1} = \frac{B_{12}\rho(\nu)}{A_{21} + B_{21}\rho(\nu)} \quad (8)$$

Si el cuerpo negro está en equilibrio térmico, la población de los niveles puede nuevamente ser calculada usando la estadística de Boltzman

$$\frac{N_2}{N_1} = \frac{g_2}{g_1} \exp\left[\frac{-h\nu}{kT}\right] \quad (9)$$

Donde g_1 y g_2 representa la degeneración (multiplicidad) de los niveles 1 y 2 respectivamente.

Fig. 4



Estamos tomando en cuenta que los niveles atómicos del oscilador pueden estar formados por varios subniveles, todos ellos muy próximos en energía.
de (8) y (9)

$$\frac{B_{12}\rho(\nu)}{A_{21} + B_{21}\rho(\nu)} = \frac{g_2}{g_1} \exp\left[-\frac{h\nu}{kT}\right] \quad (10)$$

Resolviendo para $\rho(\nu)$

$$\langle A_{21} + B_{21}\rho(\nu) \rangle \frac{g_1}{g_2} \exp\left[-\frac{h\nu}{kT}\right] = B_{12}\rho(\nu) \quad (11)$$

$$\rho(\nu) = \frac{\frac{\rho_1}{\rho_2} A_{21}}{B_{12} \exp\left[-\frac{h\nu}{kT}\right] - \frac{g_2}{g_1} B_{21}} \quad (12)$$

$$\rho(\nu) = \frac{A_{21}/B_{21}}{\left(\frac{g_1}{g_2}\right)\left(\frac{B_{12}}{B_{21}}\right) \exp\left[-\frac{h\nu}{kT}\right] - 1} \quad (13)$$

Esta ecuación es idéntica a la fórmula de Plank

$$\rho(\nu) = \frac{8\pi r^2}{c^3} \frac{h\nu}{\exp\left[-\frac{h\nu}{kT}\right] - 1}$$

sólo si

$$\frac{g_1 B_{12}}{g_2 B_{21}} = 1 \Leftrightarrow g_1 B_{12} = g_2 B_{21} \quad (14)$$

y

$$\frac{A_{21}}{B_{21}} = 8\pi \frac{h\nu^3}{c^3} \quad (15)$$

Los coeficientes están interrelacionados, por ejemplo una medida en un experimento de absorción, puede dar información del coeficiente de emisión estimulada. Ahora bien, los átomos no solamente interaccionan con radiación, también lo hacen con otros átomos y partículas. Veremos como estos procesos afectan el tiempo de vida de los niveles.

Cuando un átomo decae desde un nivel excitado de energía E_2 a otro de energía E_1 puede hacerlo emitiendo un fotón de energía

$$h\nu_0 = E_2 - E_1 \tag{16}$$

Si los valores de energía E_2 y E_1 están perfectamente definidos, la frecuencia de la luz emitida por la transición atómica sería como una delta $\delta(\nu - \nu_0)$, esto es, una línea fina como es ilustrada en la siguiente figura



Fig. 5

Sin embargo, este no es el caso y si medimos experimentalmente el ancho espectral de esa transición atómica $\Delta\nu$, tiene un valor finito, el cual es función de muchas variables. Ellas incluyen la estructura misma del átomo y por ejemplo, la temperatura y presión si estamos trabajando con un gas.

Una razón para que las líneas espectrales tengan un ancho finito, es explicada por la teoría de la Mecánica Cuántica. Ella nos dice que los valores de energía de los niveles no están perfectamente definidos, hay una “incerteza” en la medida de cada nivel. De acuerdo con (16) una incerteza en la definición del valor de los niveles de energía, se traslada a una incerteza en el ancho de la línea. Este tipo de ensanchamiento, que depende del átomo mismo y su comportamiento cuántico, es conocido como “ensanchamiento natural”.

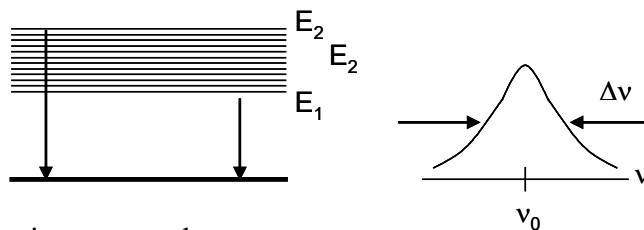
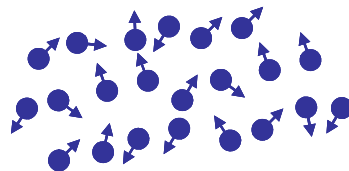


Fig. 6. Ensanchamiento natural

Hay otras razones para que el ancho de línea usualmente sea mucho más grande (más ancho) que el dado por el ensanchamiento natural. Consideremos átomos en forma de gas, como por ejemplo, N_e dentro de un tubo de un láser de He-Ne (que detallaremos más adelante). En cualquier instante de tiempo, los átomos se están moviendo en todas direcciones, con velocidades aleatorias.



Movimiento atómico en un GAS

Ahora un observador verá que los átomos que se están moviendo hacia él, emiten fotones con diferentes frecuencias que los átomos que se alejan de él, esto es conocido como el efecto Doppler. En nuestro caso, un grupo selecto de átomos emiten a una cierta frecuencia ν_0 en su sistema de referencia que se está moviendo alejándose de él en el sistema de referencia del laboratorio a una velocidad ν_z , entonces el observador medirá la frecuencia “corrida” por efecto Doppler

$$\nu'_0 = \nu_0 \left(1 - \frac{v_z}{c} \right) \quad (17)$$

donde c es la velocidad de la luz.

Dado que diferentes grupos de velocidades, emiten a diferentes frecuencias (desde el punto de vista del observador), la línea (transición), se ensancha. El resultado del ensanchamiento Doppler tiene un perfil gaussiano $g(\nu)$ dado por

$$\Delta\nu_D = \frac{[8kT \ln 2]^{1/2}}{mc^2} \nu_0 \quad (18)$$

donde k es la constante de Boltzman y T es la temperatura ambiente en grados Kelvin, m es la masa del átomo que emite. Notemos que cuanto más caliente esté el gas, mayor será la velocidad del átomo y por lo tanto, mayor será el ancho de línea.

La función $g(\nu)$ describe la distribución de intensidad emitida como función de la frecuencia ν y se la llama, “función de forma de línea”. Usando un factor de escala arbitrario tal que:

$\int_{-\infty}^{+\infty} g(\nu) d\nu = 1$, la función $g(\nu)$ puede ser interpretada como la probabilidad que el fotón que emite espontáneamente en la transición atómica, tendrá una frecuencia entre ν y $\delta\nu$.

Este tipo de ensanchamiento lo usaremos más adelante, pero además, hay otros tipos, como por ejemplo el ensanchamiento debido a las colisiones entre los átomos cuando aumenta la presión, pero estos no los consideraremos en este manuscrito.

AMPLIFICACIÓN DE RADIACIÓN

Ya vimos que los coeficientes de Einstein de emisión espontánea A_{21} , de absorción B_{12} y emisión estimulada B_{21} están relacionados (ec. 14 y 15).

$$g_1 B_{12} = g_2 B_{21}$$

Ahora usando estos conceptos, vamos a obtener las condiciones necesarias para la amplificación de luz en un material. Consideremos una onda monocromática de frecuencia ν

e intensidad $I_\nu = \rho(\nu)c = \frac{W}{m^2}$ que se propaga a través de un medio atómico con N_2 átomos por unidad de volumen en el nivel 2 y N_1 en el nivel 1

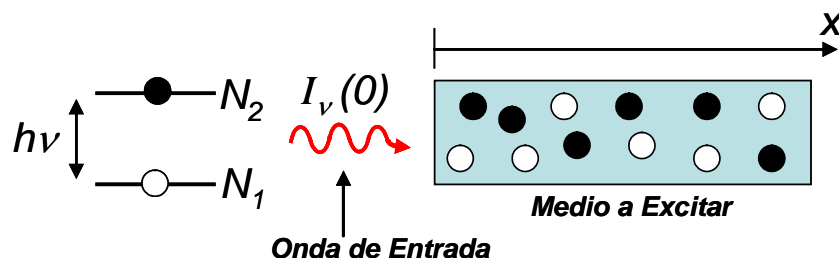


Fig. 7

La intensidad para una cierta frecuencia ν cambiará en función de la distancia como

$$\frac{dI_\nu}{dx} = N_2 B_{21} \rho(\nu) h \nu g(\nu) - N_1 B_{12} \rho(\nu) h \nu g(\nu) = g(\nu) \left(N_2 - N_1 \frac{g_2}{g_1} \right) h \nu \rho(\nu) B_{21} \quad (19)$$

pero como $B_{21} = A_{21} \frac{c^3}{8\pi\nu^3 h}$

$$\frac{dI_\nu}{dx} = g(\nu)(N_2 - N_1 \frac{g_2}{g_1}) A_{21} \frac{c^3 I_\nu}{8\pi\nu^2 c} \quad (20)$$

$$\frac{dI_\nu}{dx} = g(\nu)(N_2 - N_1 \frac{g_2}{g_1}) A_{21} \frac{c^2 I_\nu}{8\pi\nu^2} \quad (21)$$

Con una solución dada por

$$I_\nu(x) = I_\nu(0) e^{\chi(\nu)x} \quad (22)$$

donde

$$\chi(\nu) = (N_2 - N_1 \frac{g_2}{g_1}) \lambda^2 \frac{A_{21} g(\nu)}{8\pi} (cm^{-1}) \quad (23)$$

es conocida como el coeficiente de ganancia.

$$\chi(\nu) = \Delta N \sigma_{E.Esp} \quad (24)$$

$$\Delta N = N_2 - N_1 \frac{g_2}{g_1} (cm^{-3}) \quad (25)$$

donde la sección transversal de emisión estimulada está dada por

$$\sigma_{E.Esp} = (\frac{\lambda^2 A_{21}}{8\pi}) g(\nu) (cm^2) \quad (26)$$

Si está ensanchada por efecto Doppler (ec. 18) $\Delta\nu_D = \frac{[8kT \ln 2]^{1/2}}{mc^2} \nu_0$

y como $g(\nu) = \frac{1}{\Delta\nu_d}$, de (22)

$$\frac{I_\nu(x)}{I_\nu(0)} = e^{\chi(\nu)x} = G_\nu = \text{Ganancia} \quad (27)$$

Algunas veces se le llama “ganancia exponencial”.

Para tener amplificación, necesitamos tener

$$\chi(\nu) > 0 \quad (28)$$

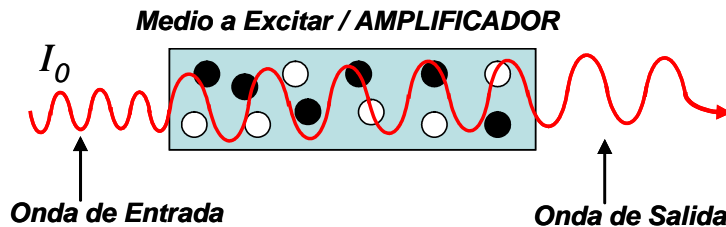
de (25) se sigue que

$$N_2 - N_1 \frac{g_2}{g_1} > 0$$

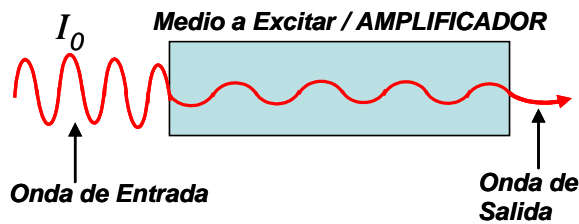
(29)

que es la condición para la amplificación.

Si $N_2 > N_1 \frac{g_2}{g_1}$ tenemos amplificación



Si $N_2 < N_1 \frac{g_2}{g_1}$ tenemos absorción



En equilibrio térmico la población de los estados deben de seguir la distribución de Boltzman, entonces

$$\frac{N_2}{N_1} = \frac{g_2 e^{-\frac{h\nu}{kT}}}{g_1} \quad (30)$$

$$\frac{N_2}{N_1 \frac{g_2}{g_1}} = e^{-\frac{h\nu}{kT}} < 1 \quad (31)$$

Por lo que, los sistemas que están en equilibrio térmico son absorbentes.

Para amplificar luz (las dos primeras letras de la palabra láser), necesitamos usar sistemas que NO ESTEN en equilibrio termodinámico. La condición puede ser satisfecha si hay una INVERSIÓN DE POBLACIÓN con respecto al equilibrio termodinámico. A menudo se refiere como “Temperatura Negativa” en la ecuación. 31.

*Notemos en (23) que el coeficiente de ganancia es directamente proporcional a la densidad de inversión de población. La población de estado superior N_2 debe de ser lo más alta posible y la del estado N_1 lo más baja posible.

**Notemos también de (23) que el coeficiente de ganancia es proporcional a λ^2 , esto significa que es más fácil obtener altas ganancias en el infrarrojo y es mucho más difícil en el

ultravioleta. Por eso, el antecesor del LÁSER fue el MÁSER donde se amplificaban microondas.

Para amplificar luz de una frecuencia ν debemos crear una inversión de población entre dos niveles de un sistema atómico separado por una energía $h\nu$

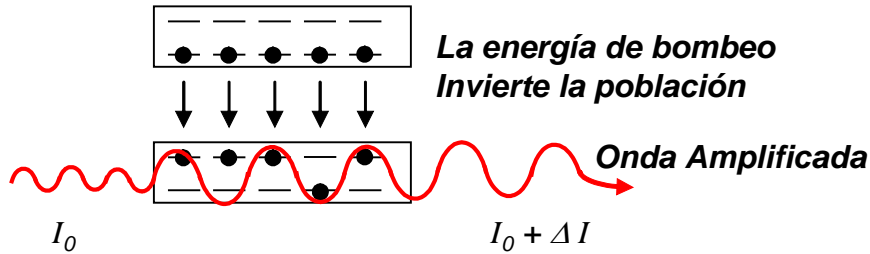


Fig. 10

En un oscilador láser, el medio amplificador tiene anisotropías, cuando lo encendemos, la emisión espontánea (ruido) comienza a ser amplificada en una dirección preferencial.

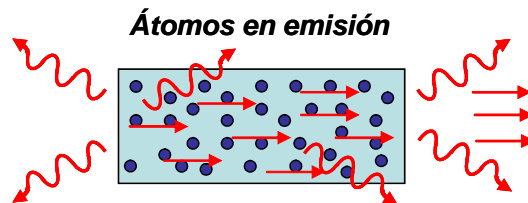


Fig. 11

Para obtener oscilación, las ganancias deben de ser mayores que las pérdidas. Para reducirlas, podemos usar resonadores ópticos (espejos), que reinyecten esta emisión a la cavidad del láser. Cualquier oscilador, es un amplificador con retroalimentación.

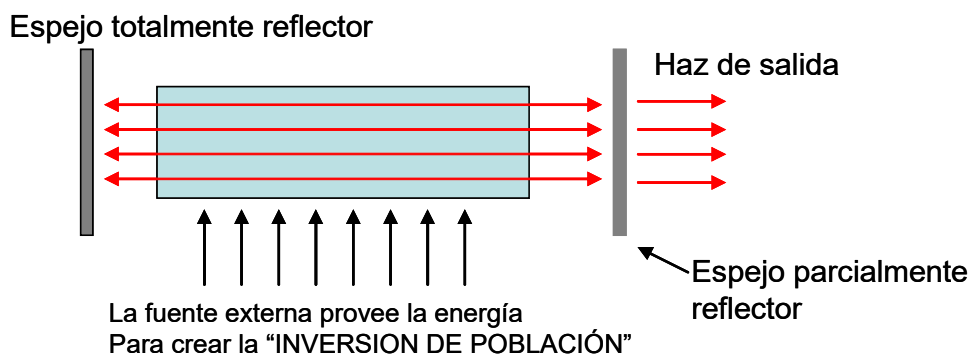


Fig. 12

Este es un esquema simplificado, que representa a la mayoría de los láseres.

EFICIENCIA CUÁNTICA

La eficiencia cuántica atómica es definida como el cociente de la energía $E_2 - E_1 = h\nu$ del fotón del láser y la energía necesaria para excitar el átomo al nivel superior E_2 . Para un sistema de 3 niveles

$$q_{ef} = \frac{E_2 - E_1}{E_3} = \frac{h\nu}{E_3}$$

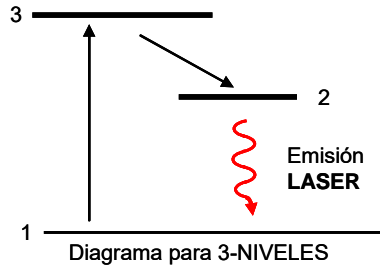


Fig.13

y para cuatro niveles

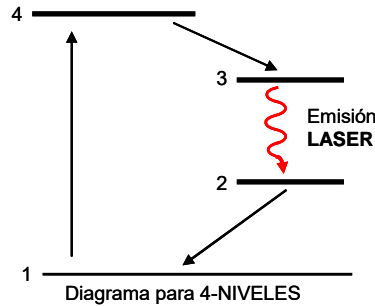


Fig. 14

$$q_{ef} = \frac{E_3 - E_2}{E_4} = \frac{h\nu}{E_4}$$

La eficiencia cuántica es la máxima eficiencia que puede tener un láser, en la práctica hay muchos factores que la limitan a valores más bajos. Como un ejemplo citemos al láser de He-Ne cuya eficiencia cuántica operando en 632.8 nm es $q_{ef} \approx 10\%$ sin embargo en la práctica, estos aparatos operan con una ganancia alrededor de 4×10^{-4} .

Los láseres de semiconductor (que no analizaremos en este manuscrito) tienen una eficiencia teórica $q_{ef} > 80\%$ y en la práctica a menudo, se logran estas eficiencias tan elevadas.

FUENTES LASER

Es posible obtener inversión de población y hacer fuentes láser usando muchos tipos de materiales.

MATERIAL	LÁSER	FUENTE DE EXCITACIÓN
sólido	Rubí, Nd:YAG	Lámpara de flash, láseres
gas (plasma)	He-Ne, ión de Ar, CO ₂	Descarga gaseosa
líquido	colorantes	Lámpara de flash y láseres
semiconductor	A _I -G _a -A _S	Fuentes de corriente, láseres

Describiremos brevemente en la siguiente tabla las características y mecanismos de excitación de los tres principales láseres gaseosos.

	λ longitud de onda en [nm]	Elemento activo	Potencia continua [W]	Excitación
He-Ne	632.8	Átomo de Ne	$1-100 \times 10^{-3}$	5 mA, 2 kV
Ar ⁺	488, 514.5, 496.5, 476.5	Ión de Ar	1-25	50 A, 500 V
Ar ⁺⁺	350	Doble ión de Ar	Hasta 4	50 A, 500 V
CO ₂	10.6×10^3	Molécula de CO ₂	10 – 1000	50 mA, 20 kV (para 100W)

Actualmente hay muchos tipos de láseres, continuos, pulsados, ultravioletas, visibles, infrarrojos, de rayos X, etc. En lo que sigue, nos centraremos en el funcionamiento de sólo dos tipos de láseres que usaremos en el experimento de laboratorio: el gaseoso de He-Ne (rojo, continuo) y el sólido de Nd:YAG (pulsado, infrarrojo y verde).

LÁSER DE HELIO NEÓN

El láser de He-Ne fue el primer láser gaseoso en operar en 1960.

Ya sabemos que debemos de excitar al medio activo, que como en este caso es un gas, simplemente con una descarga de alta tensión del gas contenido en un tubo, lo podemos lograr. Pensaríamos en un esquema como el de la figura siguiente:

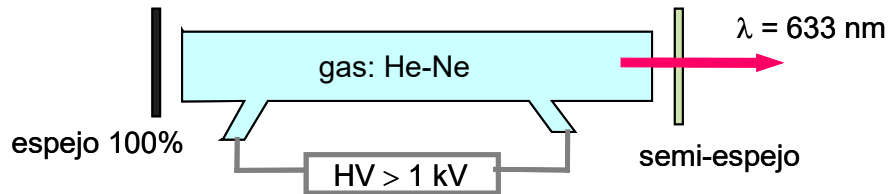


Fig. 15

Sin embargo, el diagrama real se muestra en la siguiente figura:

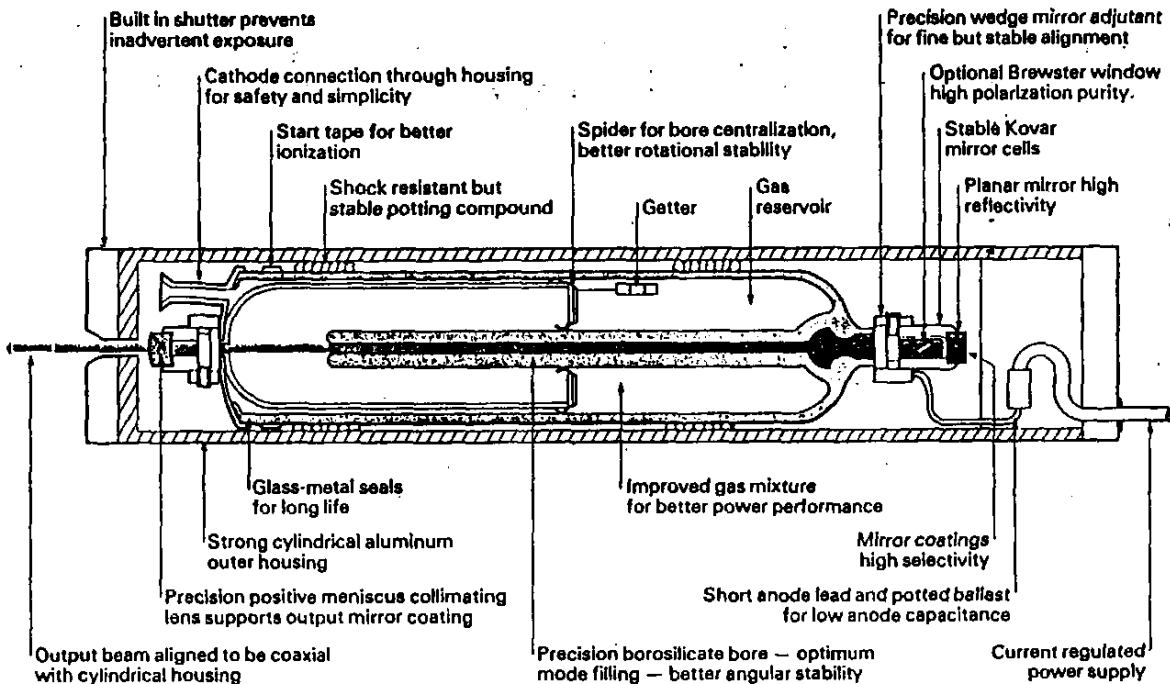


Fig. 16. Vista de la sección transversal de un láser de He-Ne comercial (Melles Griot), donde se muestran los detalles del tubo de plasma.

Como vemos, nos es simplemente una descarga gaseosa con dos electrodos, tiene mucha tecnología incorporada y esto es debido a la baja eficiencia que presenta.

El medio activo o medio de ganancia está compuesto por átomos de Neón mezclados con Helio en relaciones típicas de 1:5 a 1:20, a presiones determinadas principalmente por el diámetro (d) del tubo. La presión del gas (P) para una operación óptima, es descrita por la relación empírica $P \cdot d \cong 0.36 \text{Torr} \cdot \text{cm}$

(recordemos que la presión atmosférica sobre el nivel del mar es de 760 Torr).

La longitud del tubo es típicamente entre 10 – 100 cm y el diámetro de 0.2 a 0.8 cm.

Cuando un alto voltaje es aplicado entre los electrodos se forma una descarga gaseosa en el tubo que emite luz para todos lados. Este tipo de descarga se las llama “descargas de resplandor” (*glow discharge*). Los electrones son acelerados desde el cátodo hacia el ánodo por un campo eléctrico creado por una fuente de poder externa. En su camino, los electrones colisionan con los átomos del gas y los excitan, es decir, los átomos van a niveles de energía excitados.

Los principales conceptos relacionados con la operación del láser de He-Ne serán enunciados a continuación e ilustrados por un diagrama de niveles de energía de la figura 17.

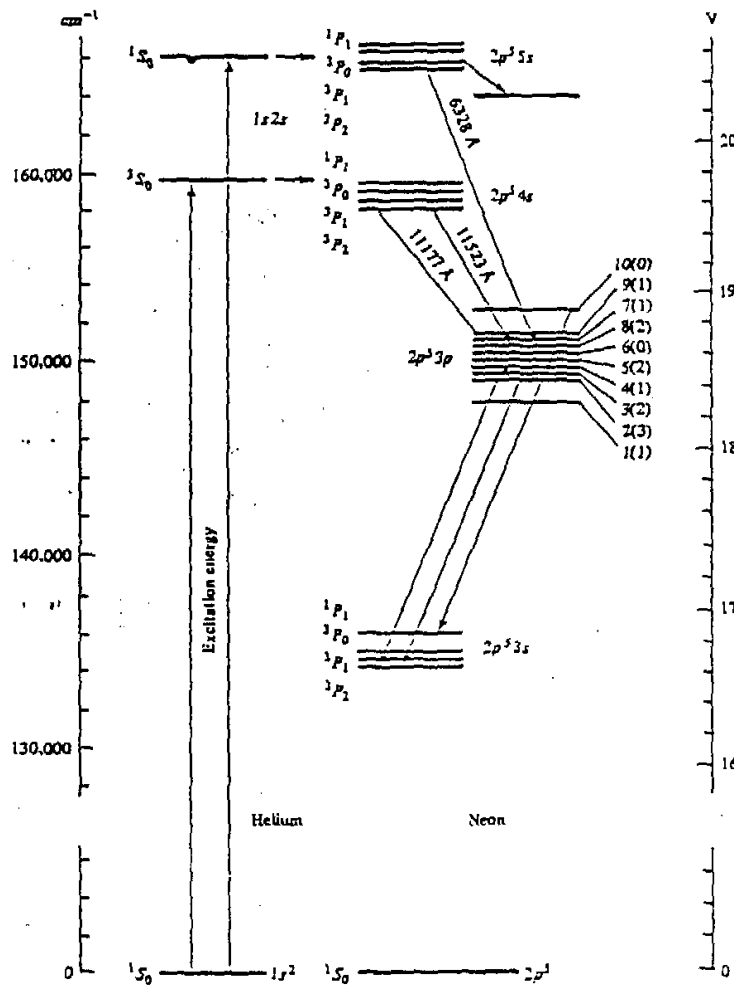
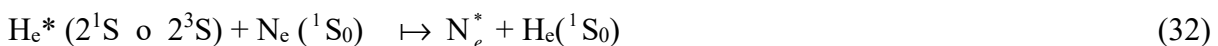


Fig. 17. Diagrama de los niveles electrónicos de energía para el He y del Ne que están envueltos en las transiciones del láser de He-Ne. (Note que las energías están medidas desde el estado fundamental).

I. El elemento activo en un láser de He-Ne es el Neón, esto es, la radiación láser emitida resulta de radiación estimulada emitida por transiciones electrónicas entre los niveles excitados de átomos de Ne. La longitud de onda más común es la de 632.8 nm, pero los hay infrarrojos (1100 nm) y de otros colores visibles.

II. En todos los casos los niveles superiores de las transiciones láser son poblados por colisiones con átomos de Helio previamente excitados a niveles de larga vida (metaestables) por colisión con electrones.



Las colisiones que excitan los niveles láser superiores del Neón son del tipo donde $N_e (^1S_0)$ y $He(^1S_0)$ son los estados fundamentales del Ne y del He, N_e^* simboliza los niveles excitados del Ne los cuales original la transición láser (ver figura 17).

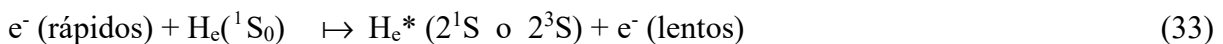
Notemos que la reacción (32) excita selectivamente los niveles láser superiores sin población directa a los niveles láser inferiores y consecuentemente, contribuyen a crear la inversión de población requerida para tener ganancia óptica.

¡Es ahora cuando entendemos la necesidad de agregarle Helio a la mezcla!!!!

Si no lo hiciéramos y tuviéramos sólo Ne, nunca lograríamos inversión de población ya que los niveles superiores e inferiores tendrían la misma población.

Los niveles láser inferiores son despopulados por decaimiento radiativo espontáneo a niveles de energía más bajos del Ne.

*Vemos que la excitación de los átomos de Helio desde su estado fundamental se da por impacto electrónico directo



Notemos que estos electrones rápidos necesitan tener al menos 20 eV. Como los niveles del He excitado $He^* (2^1S \text{ o } 2^3S)$ son metaestables, hay una gran probabilidad de que ellos colisionen con el estado fundamental de los átomos de Ne antes de que decaigan al estado fundamental del He.

** La excitación de los átomos de Neón es debida a colisiones con átomos de Helio excitados (He^*), ver ecuación 33.

Notamos de la figura 17 que lo selectivo de esta reacción, viene de la pequeña diferencia de energía entre los niveles metaestables del He y los niveles láser del Ne.

III. Contrariamente a lo que uno podría pensar en un primer análisis, incrementando la corriente de excitación arriba de cierto valor (6-20 mA dependiendo del diámetro del tubo de descarga), la potencia de salida del láser de He-Ne NO AUMENTA!!!! La razón es que, incrementando la corriente de descarga, se incrementa la producción de los niveles metaestables del He por la reacción (33), pero estos estados excitados se destruyen en una fracción importante por impacto electrónico. Esto hace que la potencia de salida del láser se sature y después decrezca como función de la corriente de excitación, cuando se sobrepasa el valor óptimo de la corriente. Esto se puede comprobar fácilmente en el laboratorio.

El láser de He-Ne es de baja potencia, si se requieren altas potencias continuas, el láser de Argón puede proveer potencias de algunos Watts.

LÁSERES DE TRES Y CUATRO NIVELES

Los láseres son clasificados comúnmente como de “tres” o de “cuatro” niveles, esto es por supuesto una simplificación, porque en cualquier láser práctico se involucran muchos niveles en el mecanismo de excitación. Lo de 3 o 4 niveles es sólo una simplificación de los niveles más significativos en los procesos de “bombeo” (excitación) o laseado. La clasificación da una idea acerca de lo fácil o difícil de bombear un medio activo para hacerlo “lasear”.

Un láser de tres niveles es aquél en el cual el nivel láser inferior es o bien el nivel fundamental o un nivel cuya separación en energía E desde el nivel fundamental es pequeña comparada con kT. Entonces en equilibrio térmico, una fracción sustancial de la población ocupa este nivel inferior de acuerdo con la relación de Boltzman.

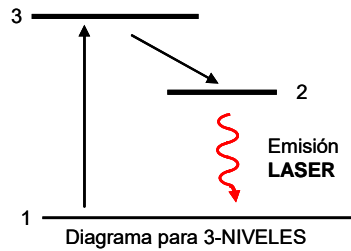


Fig. 18.

En un láser de tres niveles, es necesaria una fuente de bombeo para crear una población

$$N_2 = N_1 = \frac{N}{2} \quad (34)$$

la cuál sólo alcanza una ganancia óptica $\chi \propto (N_2 - N_1) = 0$ ¡!!!!

En la ecuación 34 asumimos que $N_3 \approx 0$ debido a las transiciones rápidas de 3 a 2, entonces $N_1 + N_2 = N$, donde N es la densidad de átomos activos.

Es decir, necesitamos una cantidad de energía para tener ganancia cero!!!!!!

Para satisfacer la condición de oscilación, el ritmo de bombeo debe de ser incrementado hasta

$$N_2 - N_1 = N_t \quad (35)$$

donde N_t aquí representa la inversión necesaria para el umbral

$$\Rightarrow N_2 = \frac{N + N_t}{2} \quad (36)$$

para sistemas de 3 niveles. La ecuación 36 nos dice la población necesaria en N_2 para lograr oscilación en el umbral.

En el caso de sistemas de “4 niveles”, tenemos una población despreciable en el nivel 1 en el caso de equilibrio termodinámico y además, si el tiempo de vida $\Gamma_1 \ll \Gamma_2$, podemos despreciar N_1 . Entonces la condición umbral es satisfecha cuando

$$N_2 = N_t \quad (37)$$

La oscilación láser comienza cuando el nivel láser superior adquiere una población igual a la del umbral N_t vía el bombeo.

La cantidad mínima de densidad de potencia necesaria para lograr y mantener la oscilación en el umbral en un sistema de tres niveles ($\frac{N_t + N_0}{2}$ átomos en el nivel superior) es:

$$P_{3niveles} = \left[\frac{N_t + N}{2} \right] \frac{h\nu}{T_2}$$

Como en la mayoría de los láseres $N \gg N_t$

$$P_{3niveles} = \frac{Nh\nu}{2\Gamma_2} \quad (38)$$

Para un sistema de 4 niveles

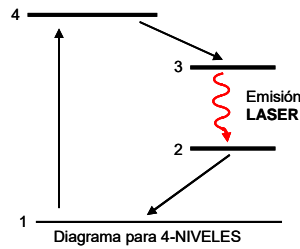


Fig. 19

La potencia requerida para el umbral

$$P_{4niveles} = \frac{N_t h \nu}{T_2} \quad (39)$$

Comparando las ecuaciones 38 y 39 obtenemos

$$\frac{P_{3niveles}}{P_{4niveles}} \approx \frac{N}{2N_t} \quad (40)$$

Para el caso de láser de Rubí, que fue el primer láser en operar en 1960, construido por T.H. Maiman en los laboratorios Hughes en EEUU y que es un sistema de tres niveles como veremos más adelante,

$$\frac{N}{2N_t} \approx 100.$$

La densidad de potencia dada por (38) y (39) es conocida como la “fluorescencia crítica”. Si el nivel 2 sólo decae radiativamente, en el umbral emitirá a través de fluorescencia una potencia P .

La ecuación (40) muestra que un sistema de cuatro niveles requiere bastante menos bombeo para crear una inversión, debido a que la población del estado más bajo (estado 1) es mínima.

LÁSERES SÓLIDOS

Ahora discutiremos láseres en los cuales el medio de ganancia está hecho de un material sólido. En estos casos el material láser es una impureza que es agregada a un material cristalino o amorfo, ejemplo de tales materiales sólidos es el Rubí y vidrios dopados.

El rubí es simplemente zafiro (Al_2O_3), un huésped cristalino básicamente transparente, dopado con pequeñas cantidades de Cromo (típicamente 5%). Los vidrios también han sido usados con éxito como huéspedes amorfos para especies láser, por ejemplo el Neodimio dopado en vidrios es una fuente láser poderosa.

La excitación de estas fuentes láser se hace usando luz, esto es una lámpara de flash u otro láser. A este tipo de excitación, se le llama “bombeo óptico”.

A continuación, discutiremos el láser de Rubí y el de Neodimio con algún detalle.

LÁSER DE RUBÍ

Este láser fue el primero en operar a mediados de 1960, en una transición a 694.3 nm correspondiente al C_r^{+3} (0.05% en Al_2O_3). Un diagrama simplificado de los niveles de energía del C_r^{+3} es mostrado en la siguiente figura.

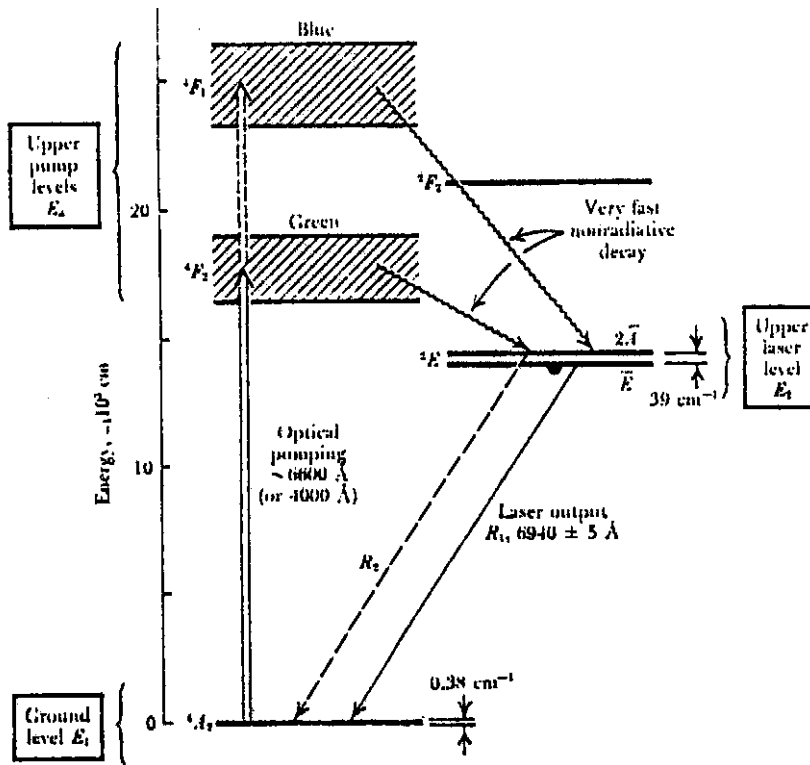


Fig. 20. Diagrama simplificado de los niveles de energía del Cr^{+3} . Láser de 3 niveles.

El tiempo de vida del nivel 3 es muy corto ($T_3 = 5 \times 10^{-6}s$) comparado con el del nivel 2 ($T_2 = 4$ ms).

Puede notarse que la mayoría de la población se origina en el estado fundamental del ión Cr^{+3} el que resulta ser el nivel inferior. Consecuentemente es un sistema de tres niveles es necesaria una considerable alta potencia de bombeo para despoblar el nivel láser inferior y lograr inversión de población y ganancia óptica.

Un esquema simplificado de la estructura de un láser de Rubí es mostrado en la siguiente figura.

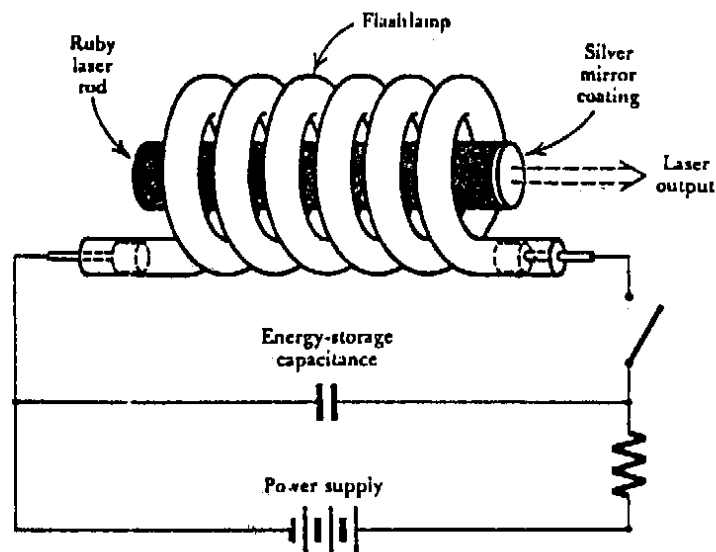


Fig. 21. Esquema simplificado de un láser de Rubí.

La barra de Rubí tiene un diámetro de 0.5 a 1 cm y una longitud de 2 a 10 cm, los extremos son paralelos y pulidos, donde se depositan los espejos directamente sobre la barra que conforma la cavidad óptica. Uno de ellos es hecho casi totalmente reflector y el otro (el de salida) es hecho parcialmente transmisor.

Una lámpara de flash de Xenón en forma helicoidal produce el bombeo óptico. La lámpara es excitada, descargando un capacitor en aproximadamente 1 ms. Esta lámpara de flash es usualmente rodeada por espejos que reflejan la radiación que se iría hacia fuera, la cuál es reflejada hacia el cristal para aumentar la eficiencia. Una buena idea es usar una lámpara lineal ubicada en el foco de una cavidad elíptica, en el otro foco, se ubica la barra de Rubí, de forma de poder bombear uniformemente el medio activo. La lámpara de flash o la barra, deben de ser rodeadas por un camisa por donde circula agua para propósitos de enfriamiento.

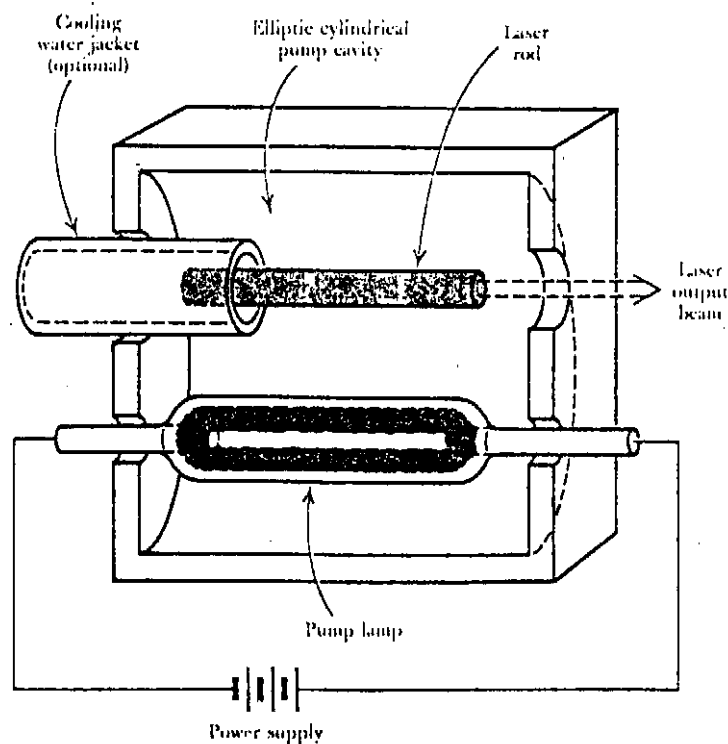


Fig. 22. Esquema para la excitación y el bombeo de una barra sólida en una cavidad elíptica.

La radiación emitida por la lámpara de flash en las regiones espectrales de 350-450 nm y 500-600 nm es absorbida por el ión de C_r^{+3} , excitando los niveles 4F_1 y 4F_2 mostrado en el diagrama de niveles de la figura 20. El coeficiente de absorción de una barra de Rubí es ilustrado en la figura 23.

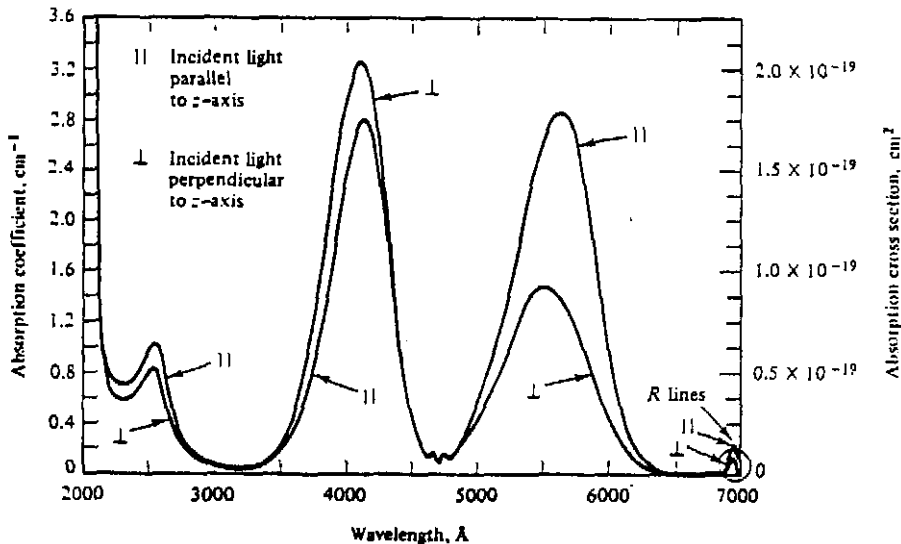


Fig. 23. Coeficiente de absorción y sección transversal de una barra de Rubí.

El mecanismo de bombeo de un láser de Rubí es el siguiente:

- 1) La radiación óptica de la lámpara de flash en la región entre 400-600 nm es absorbida por el ión C_r^{+3} en el estado fundamental, excitándolo a los estados 4F (estado 2). Esta misma radiación estimula estos átomos a retornar al nivel fundamental, pero en competencia con esta emisión estimulada, hay procesos rápidos de relajación no-radiativos del nivel 3 al nivel láser superior.

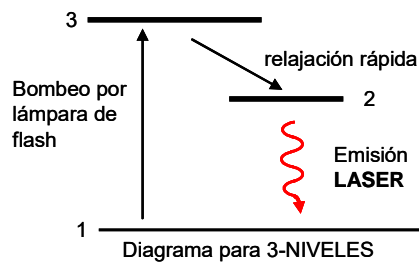


Fig. 24. Diagrama de energía simplificado del mecanismo de bombeo de un láser de Rubí.

- 2) Debido a esta relajación muy rápida del nivel 3 al nivel 2 y a la emisión espontánea muy lenta del nivel láser superior al estado fundamental ($T_{EE} = 3$ ms), los átomos son acumulados en el nivel 2, es como “un cuello de botella”, donde lo átomos quedan atrapados en este nivel. Si el proceso $1 \rightarrow 3 \rightarrow 2$ es rápido comparado con el $(T_{EE})^{-1}$, se obtiene una inversión de población entre el nivel 2 y el nivel 1.
- 3) Tendremos un láser si proveemos una retroalimentación óptica, mediante los espejos que forman la cavidad.

El pulso de la lámpara de flash, la evolución de la inversión de población y la luz de salida en 694.3 nm es mostrado en la siguiente figura.

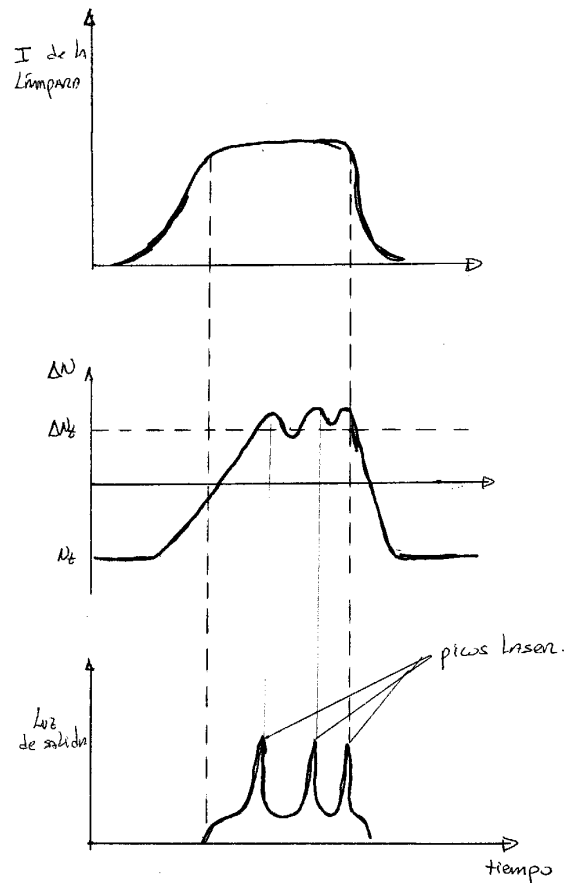


Fig. 25. Evolución de la inversión de población y la potencia de salida del láser de Rubí, en relación con el pulso de excitación de la lámpara de flash.

Puede verse que cuando la lámpara comienza a excitar a la barra, el nivel superior de la transición láser también comienza a poblarse, hasta llegar a lograr la inversión de población. En ese instante, comienza a haber emisión estimulada (láser) y se obtiene luz monocromática y coherente. Esto dura un cierto tiempo, ya que el nivel 2 se comienza a “descargar” mediante la emisión estimulada y la inversión de población decae más abajo del umbral. Como la excitación de la lámpara continúa, se vuelve a restablecer la emisión estimulada y nuevamente volvemos a tener emisión láser en forma de pulsos. Es decir que en un solo pulso de descarga de la lámpara de flash, tenemos varios pulsos láser con una duración mucho menor. Hay técnicas que permiten obtener un solo pulso, la más simple, sería que el pulso de excitación de la lámpara se acabe justo cuando se llega a la máxima inversión de población, pero hay varios otros métodos para modificar la forma del pulso láser.

Para una energía de excitación de la lámpara de 1000 J, se obtiene usualmente salida láser, en $>100 \mu\text{s}$ con potencias pico de 100 kW y energías del orden de 10 J. Esta potencia está concentrada en un haz altamente monocromático y direccional que permite detectar ecos de radar de la luna o atravesar placas de acero de 1 cm de espesor en un solo pulso.

LÁSER DE NEODIMIO

Otro material sólido muy usado como medio de ganancia óptica es hecho dopando Neodimio en una gran variedad de materiales huéspedes. Los dos más comunes son el vidrio y el Itrio Aluminio Granate ($\text{Y}_3\text{Al}_5\text{O}_7$), abreviado YAG.

La figura 26 muestra un diagrama simplificado de los niveles de energía del Nd^{+3} para vidrio y YAG como materiales huéspedes.

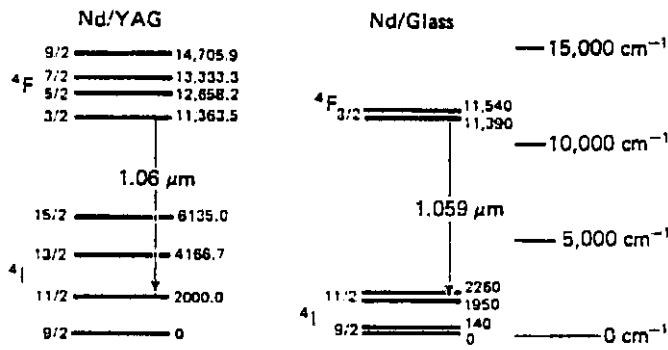


Fig. 26. Diagrama simplificado del Nd^{+3} en vidrio y YAG. Láser de 4 niveles.

El Nd^{+3} lasea en el infrarrojo cercano, aproximadamente en $1.06 \mu m$ y es un láser de 4 niveles. Consecuentemente requiere significante menos energía de bombeo que el láser de Rubí para lograr la inversión de población.

Calculemos la fracción de población del nivel láser inferior ${}^4I_{11/2}$ a temperatura ambiente $T = 300 K$, $kT_e = 208.3 \text{ cm}^{-1}$

La energía del ${}^4I_{11/2}$ es 208.3 cm^{-1}

$$\frac{N({}^4I_{11/2})}{N({}^4I_{9/2})} = \frac{g_{(11/2)}}{g_{(9/2)}} \exp\left(\frac{-\Delta E}{kT}\right) = \frac{12}{10} \left(\frac{-2000}{208.3}\right) = 8 \times 10^{-5}$$

Esta cuenta indica que la población inicial del nivel láser inferior es despreciable comparada con la población del nivel ${}^4I_{9/2}$. Es por lo tanto mucho más fácil crear y mantener una población residual bombeando al nivel ${}^4F_{3/2}$.

El espectro de transmisión del Nd:YAG se muestra en la figura 27. No es un asunto trivial lograr la correspondencia del espectro de emisión de una lámpara con un espectro tan complicado.

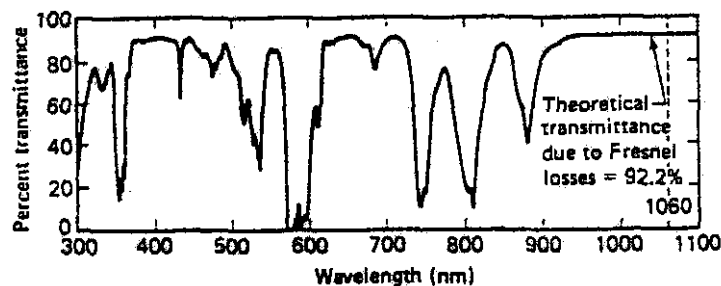


Fig. 27. Espectro de transmisión del Nd:YAG de una muestra de 6.4 mm de espesor.

De toda la energía que emite la lámpara, solo nos interesan aquellas frecuencias que sean absorbidas por la barra que bombeen los átomos de Nd al nivel láser superior, el resto lo que hace es calentar todo el sistema y es por eso que los láseres bombeados por lámparas de flash necesitan sistemas de refrigeración. Para evitarlos, podemos bombear el medio activo con un láser que emita en la región que nos interesa, este es el caso de los láseres de Nd:YAG bombeados por láseres de semiconductor, emitiendo en 800 nm.

Una de las diferencias más significativas entre el YAG y el vidrio, es el ancho de línea de la transición $1.06 \mu m$ en ambos materiales.

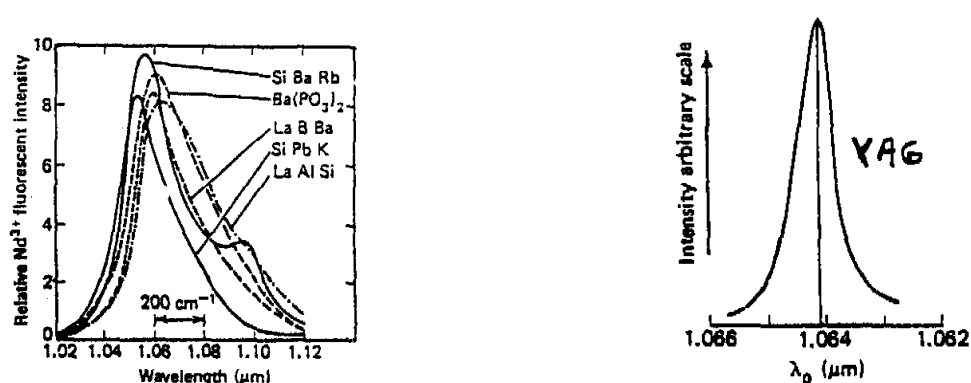


Fig. 28 Anchos de línea para la transición del Nd^{+3} en $1.06 \mu\text{m}$ en diferentes materiales y en YAG.

El ancho de línea del YAG ($\Delta\lambda_{\text{YAG}} = 0.7 \text{ nm}$) es mucho menor que para el vidrio, $\Delta\lambda_{\text{vidrio}} = 30 \text{ nm}$. La mayor diferencia entre los anchos de línea es debido al rol de los átomos en el material huésped. El YAG al ser una sustancia cristalina bien definida, presenta un ambiente cercano igual para cada ión Nd^{+3} y como consecuencia de esto, se observa un ancho de línea angosto. Por otro lado, el vidrio al ser un material amorfo, presenta un sitio diferente en cada ubicación dando consecuentemente líneas más anchas.

Como la ganancia de pequeña señal es inversamente proporcional al ancho de línea, el láser YAG tiene casi 40 veces más ganancia que para el Nd^{+3} en vidrio, o dicho de otra manera, se necesita una inversión de población 40 veces menor para alcanzar el umbral de oscilación.

Consecuentemente el láser de YAG es mucho más fácil de bombear y no es sorprendente que la mayoría de los láseres continuos en $1.06 \mu\text{m}$ sean de la variedad YAG. Sin embargo, los láseres grandes, es decir de altas potencias, como por ejemplo los que se utilizaban en fusión donde se almacenan altas energías, son láseres de vidrio. El vidrio es una sustancia que ha sido estudiada por siglos y se puede hacer muy uniforme con pulimientos muy precisos.

De todo lo anterior, hay muchos libros básicos sobre los principios y las aplicaciones de los láseres, al final se presenta un lista de algunos de ellos.

EXPERIMENTO

La idea del experimento, es mostrar a los láseres “en acción”. Se trabajará con los láseres que hemos estudiado en detalle (He-Ne y Nd:YAG), uno continuo y el otro pulsado.

El objetivo, es medir las velocidades de las ondas de choque generadas por un plasma inducido por láser en aire, mediante la técnica de deflectometría.

Los alumnos manejarán además de los láseres, componentes ópticos y mecánicos de precisión (lentes, filtros interferenciales, mesas de traslación micrométrica, etc.) y componentes y aparatos electrónicos (fotodetectores, medidores de energía osciloscopios, computadoras, etc.). Por último, los alumnos deberán de presentar una medida de un fenómeno físico.

PLASMAS INDUCIDOS POR LÁSER

El plasma es un estado de la materia, más del 90% del Universo es plasma, los sólidos, los líquidos y los gases son de hecho una singularidad. El Sol es un plasma, los tubos fluorescentes de iluminación también y si buscamos aquí en la Tierra, veremos que hay más plasmas de lo que pensamos.

Cuando la materia (que puede ser sólida líquida o gaseosa) es suficientemente excitada, algunos átomos se ionizan y esto es ya un plasma. El más simple que podemos hacer, es excitar el aire que nos rodea, por ejemplo, mediante una descarga eléctrica. Otra forma es usar luz, energía que puede ser enfocada mediante una lente. Con una lupa podemos encender el fuego concentrando la luz del Sol. Si usamos láseres, que son amplificadores de luz, podremos lograr la excitación de los átomos y moléculas que componen el aire y formar un plasma. Si la excitación es mediante un pulso de luz, provocaremos una explosión que tiene una onda de choque asociada, provocada porque la velocidad de las partículas es mayor que la velocidad del sonido en el medio. Además, el gas se calienta y se expande en forma adiabática, enfriándose al aumentar su volumen.

La formación y evolución del plasma producido por pulsos láser en aire no es un fenómeno totalmente entendido [1]. Si bien los primeros estudios datan de la fecha en que fueron accesibles los láseres pulsados [2], estos han continuado hasta el presente y debido a lo complejo del problema, es una línea de investigación que continúa abierta [3].

Los plasmas inducidos por láser tiene aplicaciones prácticas según sean producidos en sólidos (ablación láser), por ejemplo para la obtención de películas delgadas [4]; en líquidos para aplicaciones médicas y técnicas como la caracterización de materiales [5] y por último en gases, para la detección de contaminantes mediante técnicas espectroscópicas (espectroscopias de rompimiento inducidas por láser) [6], entre otras. En el Laboratorio de Fotofísica del CCADET de la UNAM trabajamos con plasmas generados en estos tres medios (sólidos, líquidos y gases) en distintas aplicaciones. Para estas, es necesario entender la física del plasma, ya que éste es el precursor de varios efectos en los que se basan las aplicaciones.

Cuando un láser pulsado de alta potencia es enfocado con una lente, si la energía es suficientemente alta, ocurre el rompimiento eléctrico del medio debido a que el campo eléctrico del haz enfocado, excede el umbral de ruptura dieléctrica del mismo; este rompimiento puede atribuirse a dos mecanismos: la ionización multifotónica directa e ionización por impacto directo o cascada. Para gases [7], el tipo de mecanismo predominante, la evolución temporal, la emisión óptica, los fenómenos microscópicos y macroscópicos del plasma, dependen de varios parámetros como son: la longitud de onda de láser (UV, VIS o IR), el ancho temporal del pulso (de los fs a los μ s), la energía, la distribución espacial de la intensidad, el tipo de gas y su presión. En este experimento analizaremos el plasma formado por pulsos de luz láser visible (532 nm) de 7 ns de duración, en aire a presión atmosférica y su posterior expansión.

La técnica que se utilizará, es simple, sencilla y económica y será la de deflectometría láser (*laser deflection or photoacoustic probe beam deflection*) [8]. Básicamente consiste en que, cuando un haz láser continuo atraviesa un medio que sufre un cambio de índice de refracción en el tiempo (formación y expansión del plasma), el haz experimenta una desviación que puede ser medida simplemente con un fotodiodo/fotodetector. Esta técnica, a pesar de su simplicidad, ha permitido obtener información importante sobre las formas y velocidades de las ondas de choque, de las ondas térmicas y de los tiempos de relajación del gas en el punto focal de la lente. Estos plasmas inducidos por láser, sus aplicaciones y las técnicas de diagnóstico (incluida la deflectometría) han sido trabajados en el Laboratorio de Fotofísica del CCADET de la UNAM en los últimos años [9-24]. Una revisión de las aplicaciones de estos plasmas en la simulación de relámpagos atmosféricos, interacción de explosiones y jets astrofísicos puede encontrarse en la referencia [12].

ARREGLO EXPERIMENTAL

El montaje experimental se muestra en la figura 29. El plasma es producido por un láser Nd:YAG pulsado marca Quantel, modelo Brilliant B que emite su línea fundamental en 1064 nm.

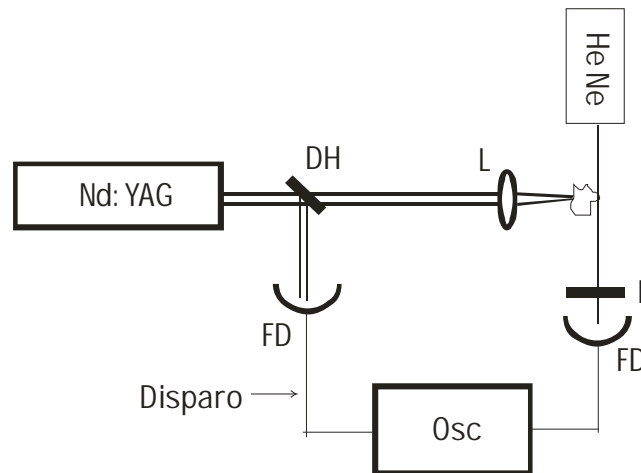


Fig.29. Arreglo experimental para el diagnóstico de una onda de choque por deflectometría. DH: divisor de haz, L: lente, FD: Fotodetector, F: filtro interferencial, Osc: osciloscopio.

Para este experimento se empleó el segundo armónico (la mitad de la longitud de onda), es decir la línea de 532 nm. Una de las ventajas de los láseres es su monocromaticidad, pero por otro lado, es como tener una radio que funciona en una sola frecuencia. Si queremos que el mismo láser opere en otras longitudes de onda, podemos usar procesos no lineales (no cubiertos en este manuscrito) y en el laboratorio, mostraremos la generación del segundo armónico: un haz infrarrojo es transformado en uno verde. Hay otras formas de cambiar la longitud de onda de emisión de los láseres, inclusive, hacerlos sintonizables (como una radio), pero esto tampoco lo cubriremos y simplemente mostraremos un láser, que como una radio, lo podemos sintonizar en distintas frecuencias, es decir, lo podemos hacer emitir en varios colores.

El ancho temporal del pulso es de 7 ns y se puede operar hasta una frecuencia de repetición de 10 Hz, es decir, podemos tener un pulso cada 100 ms. El haz se enfoca en el aire mediante una lente plano-convexa de 5 cm de distancia focal con un recubrimiento antirreflejante (KPX082AR.33 de Newport). La lente se monta sobre una mesa micrométrica con el fin de poder variar la posición de la formación del plasma.

El haz de prueba es un láser He-Ne de 1 mW de potencia emitiendo a 632.8 nm (modelo U-1301 de Newport). La deflexión del haz se detecta por un fotodiodo, el cual enviaba la señal a un osciloscopio (TDS 524A de Tektronix) disparado con el pulso del láser Nd:YAG. Para evitar el ruido proveniente de luz de otras longitudes de onda se coloca un filtro interferencial en el fotodiodo que recibe la luz del láser He-Ne. La adquisición de datos se debe de realizar haciendo promedio de señales en el osciloscopio digital.

El profesor guiará las actividades para estas medidas, teniendo muy en cuenta las indicaciones de seguridad, ya que se trabajará con fuentes de alta tensión, sustancias cancerígenas y haces intensos de luz.

AGRADECIMIENTOS

A la Dirección General de Asuntos del Personal Académico (DGAPA) de la UNAM y al Consejo Nacional de Ciencia y Tecnología (CONACyT), México y a la Facultad de Ciencias de la UNAM.

REFERENCIAS BÁSICAS

- “Laser fundamentals” W. Silfvast, Cambridge University Press, (1996).
- “Tunable laser applicatios” F. Duarte, Marcel Dekker Inc., (1995).

- "Laser electronics" J.T. Verdeyen, Prentice-Hall, (1989).
- "Lasers and optical engineering" P. Das, Springer-Verlag, (1990).
- "Industrial laser and their applications" J. Luxon and D. Parker, Prentice-Hall, (1992).
- "Laser tecnology" S. Lugomer, Prentice-Hall, (1990).
- "Encyclopedia of laser and optical technology" A. Meyers, Academic Press, (1991).
- "Solid state laser engineering" W. Koechner, Springer-Verlag, (1992).
- "The laser guide book" J. Hecht, McGraw-Hill, (1986).
- "The industrial laser handbook" D. Bolforte and M. Levitt, Sringer-Verlag, (1992).
- "Láseres Sintonizables de Estado Sólido y Aplicaciones", J. Fernández, et al., Ediciones de la Universidad Autónoma de Madrid, España, (1989).

REFERENCIAS PARA EL EXPERIEMNTO

1. J. F. Kielkopf, *Physical Review E*, **52**, (2), 2013-2024, (1995).
2. P.K. Ready, *Efects of High-Power Laser Radiation*, Academic, London, (1971).
3. K. Ito and T. Aizawa, *J. Mat. Proc. Tech.* **85**, 91-96, (1999).
4. D.B. Chrisey and G.K. Hubler (ed.), *Pulsed laser deposition of thin films*, John Wiley, New York, (1994).
5. J. Noack, D.X. Hammer, G.D. Noojin, B.A. Pockwell and A. Vogel, *J. Appl. Phys.* **83** (12), 7488-7495, (1998).
6. K.Y.Yamamoto, D.A. Cremers, M.J. Ferris and L.E. Foster, *Appl. Spec.* **50** (2), 222-233, (1996).
7. R. Tambay and R.K. Thareja, *J. Appl. Phys.* **70** (5), 2890-2892, (1991).
8. D.P. Almond and P.M. Patel, *Photothermal science and techniques*, Chapman & Hall, London, (1996).
9. A. Robledo-Martínez, H. Sobral, M. Villagrán-Muniz, 2004, "Dipolar field and plasma expansion at the onset of laser-induced breakdown in a uniform DC field" *Journal of Physics D: Applied Physics*, en prensa, 2004.
10. C. Sánchez Aké, H. Sobral, E. Sterling, M.Villagrán-Muniz, 2004, "Time of Flight Of Dual-Laser Ablation Carbon Plasmas by Optical Emission Spectroscopy", *Applied Physics A; Material Science and Processing*, vol. 79, 1345-1347. DOI: 10.1007/s00339-004-2775-y.
11. H. Sobral, R. Sanginés de Castro, M. Villagrán Muniz, R. Navarro González, 2004, "Time Resolved Study of Simulated Volcanic Lightning by Laser Ablation of Basalt in Air", *Geophysical Research Letters*, Vol. 31, L07101, doi:10.1029/2003GL019217.
12. M. Villagrán-Muniz, H Sobral, R. Navarro-González , P. F. Velázquez and A. C. Raga, 2003, "Experimental Simulation of Lightning, Interacting Explosions and Astrophysical Jets with Pulsed Lasers", *Plasma Physics and Controlled Fusion*, Vol. 45, pp. 571-584.
13. C. Sánchez Aké, H. Sobral, M. Villagrán Muniz, L. Escobar-Alarcón y E. Camps, 2003, "Characterization of laser ablation plasmas by laser beam deflection", *Optics and Lasers in Engineering*, vol 39, no 5-6, pp. 581-588.
14. M. Villagrán-Muniz, H Sobral, R. Navarro-González, 2003, "Shock and Thermal Wave Study of Laser-Induced Plasmas in Air by Probe Beam Deflection Technique", *Measurements Science and Technology*, vol 14, pp. 614-618.
15. F. Bredice, D. Orzi, D. Schinca, H. Sobral, M.Villagrán-Muniz, 2002, "Characterization of pulsed laser generated plasma through its perturbation in an electric field", *IEEE Transactions on Plasma Science*, Vol.30, No.6, pp.2139-43.
16. H. Sobral, M. Villagrán-Muniz, R. Navarro-González, E. Camps, 2002, "Experimental simulation of a doubled return-stroke lightning flash by laser", *Geophysical Research Letters*, vol 29, No. 23, pp. 2086-2089.

17. L. Escobar-Alarcón, E. Camps, E. Haro-Poniatowski, M. Villagran, C. Sánchez, 2002, "Characterization of rear and front-side laser ablation plasmas for thin film deposition", *Applied Surface Science*, 197-198, pp. 192-196.
18. R. Navarro-González, M. Villagrán-Muniz, 2001, "Effect of beam waist on shock wave properties of laser-induced plasma in air by the photoacoustic probe beam deflection method", *Analytical Sciences*, Vol. 17, pp. s118-s121.
19. P. Velazquez, A. Raga, H. Sobral, M. Villagran-Muniz, R. Navarro-González, 2001, "Laboratory experiments and numerical models of interacting explosions", *Revista Mexicana de Astronomía y Astrofísica*, Vol. 37, no 1, pp. 87-94, 2.
20. A. Raga, H. Sobral, M. Villagrán-Muniz, R. Navarro-González, E. Masciadri, 2001, "A numerical and experimental study of the development of turbulence in a low Mach number jet", *Monthly Notices of the Royal Astronomical Society*, Vol. 324, No.1, pp. 206-12, 7.
21. M. Villagrán-Muniz, H. Sobral, E. Camps, 2001, "Shadowgraphy and interferometry using a CW laser and a CCD of a laser-induced plasma in atmospheric air", *IEEE Transactions on Plasma Science*, Vol. 29, no 4, pp. 613-616, 3
22. R. Navarro-González, M. Villagrán-Muniz, H. Sobral, L. T. Molina y Mario Molina, 2001, "The physical mechanism of nitric oxide formation in simulated lightning", *Geophysical Research Letters*. Vol. 28 , No. 20 , pp. 3867-3870, 2.
23. H. Sobral, M Villagran Muniz, R. Navarro González y A. Raga, 2000, "Temporal evolution of the shock wave and hot core air in laser induced plasma", *Applied Physics Letters*, Vol. 77, no 20, pp. 3158-3160, 15.
24. A.C. Raga, R. Navarro-González, M. Villagrán-Muniz, 2000, "A new, 3D adaptive grid code for astrophysical and geophysical gasdynamics" , *RMxAAA*, Vol 36, No. 1, pp. 67-76.

Dr. Mayo Villagrán Muniz
Agosto, 2004.