

CIENCIA,

y un gran paso para la humanidad !!!

← Earth



COORDINADORES: Ana S. CASALVILLA DUEÑAS y Quintín GARRIDO GARRIDO

AUTORES:

Bernardo ADEVA ANDANY, Claudi ALSINA, Enrique ÁLVAREZ, Juan A. de CARLOS SEGOVIA, Marta DUEÑAS PORTO, Luis ESPAÑOL GONZÁLEZ, Elías FERNÁNDEZ-COMBARRO ÁLVAREZ, Leonardo FERNÁNDEZ JAMBRINA, Diego GARCÍA-MARTÍN, Claudia GARCÍA GARCÍA, José Miguel GARCÍA-MARTÍN, Antonio GÓMEZ TATO, Francisco A. GONZÁLEZ REDONDO, Ester LÁZARO LÁZARO, Emma LÓPEZ-ALONSO CONTY, Ramon LÓPEZ de MÁNTARAS BADÍA, Xabier MARCANO, Manuela MARTÍN SÁNCHEZ, Martín MARTÍNEZ-RIPOLL, Alfonso OGAYAR SERRANO, Jorge OLCINA CANTOS, Gabriel PINTO CAÑÓN, Carolina RUBIO MARTÍN, Beatriz SÁNCHEZ-CANO, Mariano SANTANDER NAVARRO, Cristian SUÁREZ CABRERA, Sonia VILLAPOL, Olivier WITASSE.

PARTICIPACIÓN ESPECIAL: José Manuel GRANDELA DURÁN con "La conquista de la Luna: España lo hizo posible"

PRÓLOGO: Michael LÓPEZ-ALEGRÍA

CIENCIA, y un gran paso para la humanidad!!!

Cubierta:

Inspiración de portada. Cielo de Marte: Ilustración sobre fotografía tomada por el Rover Curiosity el 31 de Enero de 2014. [Archivo NASA].

Montaje cortesía de Luis Bretón Belloso.

Este libro tiene una licencia Creative Commons.

Cualquier reproducción total o parcial de esta obra deberá hacer un reconocimiento expreso a la autoría de la misma y/o de los capítulos mencionados.

No se permite el uso comercial de la obra original ni de las posibles obras derivadas.

La distribución y uso de las obras derivadas se debe hacer bajo una licencia igual a la que se regula la obra original.

Las imágenes, figuras, ilustraciones que aparecen están amparadas bajo esta licencia salvo en las que expresamente se hace mención a su autoría/crédito al pie de las mismas.



Presentación y agradecimientos.

“Un gran paso para la Humanidad”, este conjunto de palabras puestas en un título hacen presagiar algo fascinante y no merecen menos que el libro que las recoge sea “Un gran libro”.

Este libro que estáis comenzando a leer pretende ser un pequeño lienzo en el que extraordinarios “artistas” han reflejado con trazos muy precisos grandes pasos científicos que han sido o serán determinantes para la Humanidad.

Todo ha sido posible por la participación de esta serie de “artistas”, gracias a su altruista colaboración este proyecto se ha podido llevar a cabo. Y sí, les consideramos “artistas” ya que no solo son unos excelentes científicos sino que además tienen la habilidad de contarnos, de hacernos partícipes, de esta querida Ciencia que para muchos de nosotros es de difícil comprensión. Y lo hacen con un lenguaje sencillo, con una pasión contagiosa, al alcance de todos los públicos e invitando al lector a ser parte principal e imprescindible de la Historia de la Ciencia.

Queremos agradecer a todos los autores participantes su colaboración y paciencia para con estos “coordinadores” que hace un año les asaltamos con esta “idea peregrina” y que nos han permitido tener el honor y el placer de compartir estas líneas con ellos.

Especial mención y agradecimiento para José Manuel, su participación era crucial para poder desarrollar este proyecto. No podemos dejar de agradecer a Michael su prólogo y sobre todo sus “grandes pasos (paseos, EVA) espaciales” que todos consideramos que también son nuestros. Gracias a Luis por su trabajo con la portada y por la paciencia que ha mostrado hasta que la hemos acabado.

Este *CIENCIA, y un gran paso para la humanidad!!!* culmina la “trilogía”, aunque solo sea por el comienzo de los títulos, que se inició con *CIENCIA, y además lo entiendo!!!* y que siguió con *CIENCIA, y yo quiero ser científico!!!*, destacar que prácticamente dos centenas de científicos han participado en estos libros (en los tres de manera altruista lo que ha permitido que los tres sean de descarga, en pdf, gratuita), todos con enorme humildad y con extraordinaria vocación de hacer y difundir buena ciencia. A todos ellos junto a otros muchos que pese a no haber podido participar por diversos motivos nos han infundido grandes dosis de ánimo y apoyo, a todos ellos les damos nuestra más sincera gratitud por su colaboración y amistad.

Ni que decir tiene que no es casual que en este 2019 se cumplan 50 años de que el hombre pisara la Luna por primera vez. El germen de este proyecto, de esta idea peregrina, de este libro, fue un regalo de cumpleaños que era una maqueta del Saturno V. Durante los fines de semana de construcción de esta maqueta se fue fraguando la idea de este libro, siempre teniendo en mente que tendría que estar listo para antes del mes de Julio. Y por cierto al “constructor de maquetas” también le caen los 50 en este Julio de 2019, considerar este libro una especie de “regalo” que este “niño de ideas peregrinas” os hace a todos con la esperanza de que lo disfrutéis y que sigamos a base de “¿y por qué?” dando “Grandes pasos para la Humanidad”.

Para terminar, lo más importante, gracias a nuestras familias por el apoyo, ánimo y comprensión sobre todo al utilizar multitud de “momentos familiares” para poder sacar adelante este libro.

Y por supuesto, compartir y difundir que para eso lo hemos hecho.

Madrid, Abril – 2019.

Ana S. Casavilla Dueñas

Quintín Garrido Garrido

Prólogo.

CIENCIA. Sin lugar a dudas, es una palabra tan profunda como intimidante. Nadie puede captar la totalidad de su significado. Y algunos dirían que, para la mayoría de nosotros, comprenderla, incluso mínimamente, no tiene mayor propósito que satisfacer un anhelo. Aunque afecta a cada segundo de nuestras vidas, lo hace sin reconocimiento ni comprensión. Como la gravedad, una de sus manifestaciones más ubicuas, está ahí. Tanto si nos guste, la conozcamos, la entendamos – o no – simplemente. . . es.

Sin embargo, el conocimiento de la Ciencia es absolutamente imprescindible para el desarrollo continuo de nuestra especie. Así como la comida y el refugio forman la base de la jerarquía de necesidades de Maslow, la Ciencia es el pilar sobre el cual la humanidad evoluciona su interacción con nuestro entorno. Sin él, no hay ingeniería ni diseño, ni fabricación ni construcción, ni seguridad ni comodidad, ni eficiencia ni innovación. Al fin y al cabo, sin entender la Ciencia, no hay progreso.

Aquellos de nosotros que queremos satisfacer el anhelo no podemos simplemente recurrir a los expertos que entienden profundamente los diversos ladrillos que forman el muro de las disciplinas científicas que sustentan nuestra comprensión de nuestro mundo. Confiamos en aquellos pocos elegidos que no solo entienden estas disciplinas, sino que también pueden explicarlas a aquellos de nosotros que no hemos pasado nuestras vidas en investigación y experimentación. Necesitamos traductores que puedan hablar no solo los diversos dialectos de estas especialidades, sino también el lenguaje genérico de la vida cotidiana. En este volumen se reúne un grupo de estas mismas personas divulgadoras: grandes mentes, todas con reconocido prestigio en sus respectivas áreas, presentan sus temas de manera simple y amena. Lo hacen con pasión y altruismo, con el objetivo de difundir su comprensión, incluso a un nivel básico, para aquellos de nosotros que no estamos satisfechos con la mera guía del usuario.

Hablar de Ciencia es hablar de multitud de grandes frases que se han dicho y se han transmitido de generación en generación sin una certidumbre sobre la autoría. Frases como “...dadme un punto de apoyo y...”, “...y sin embargo se mueve...”, “...a hombros de gigantes...”, “...Dios no juega a los dados...” y por supuesto “That’s one small step for (a) man; one Giant leap for mankind” (Éste es un pequeño paso para el hombre; una zancada gigante para la humanidad). Esta frase y que en este año de 2019 se cumplan 50 años de que fuera dicha (y sí, hay certidumbre de que fue dicha, por quién fue dicha y dónde fue dicha) son el hilo conductor de este libro. Es un auténtico privilegio

poder disfrutar del capítulo de José Manuel Grandela, poder saber y conocer, de primera mano, una parte de aquellos grandes momentos de 1969.

Es de destacar y agradecer a Ana y Quintín el no rendirse a la tentación de haber conformado una lista con sus “grandes pasos para la humanidad”, cosa que hubiera llevado a una infructuosa discusión sobre la importancia de unos sobre otros, de los que están y de los que faltan.

Se puede decir sin temor a equivocarnos que los capítulos que componen este libro son todos “Grandes pasos para la humanidad” tanto en la parte de “Fueron...” como en la de “Serán...”. Este libro se puede entender como ese pequeño grano de arena que nos devuelve el océano del conocimiento de vez en cuando.

Michael López-Alegría
Astronauta.
Principal. MLA Space, LLC
Washington, D.C.

Índice.

| | |
|---|------------|
| Presentación..... | 5 |
| Prólogo..... (Michael López-Alegría) | 7 |
| Índice..... | 9 |
| 1. La conquista de la Luna..... (José Manuel Grandela Durán) | 11 |
| <i>1ª Parte – Fueron grandes pasos...</i> | |
| 2. La tabla periódica..... (Gabriel Pinto Cañón) | 42 |
| 3. La creación de los números..... (Claudi Alsina) | 56 |
| 4. Rayos X..... (Martín Martínez-Ripoll) | 69 |
| 5. Darwin y la evolución..... (Alfonso Ogayar Serrano) | 80 |
| 6. El cálculo infinitesimal..... (Luis Español González) | 92 |
| 7. La Electroquímica..... (Manuela Martín Sánchez) | 106 |
| 8. Vacunas..... (Carolina Rubio Martín) | 119 |
| 9. El Electromagnetismo..... (Diego García-Martín) | 127 |
| 10. En la Prehistoria de la Astronáutica..... (Francisco A. González Redondo) | 143 |
| 11. Tiempo y clima..... (Jorge Olcina Cantos) | 155 |
| 12. Santiago Ramón y Cajal..... (Juan A. de Carlos Segovia) | 171 |

| | | |
|------------|---|------------|
| 13. | La Gravedad..... | 183 |
| | (Leonardo Fernández Jambrina) | |
| 14. | Descifrando el genoma humano. | 195 |
| | (Cristian Suárez Cabrera) | |
| 15. | El Modelo Estándar de la Física de Partículas..... | 205 |
| | (Claudia García García) | |

2ª Parte – Serán grandes pasos

| | | |
|------------|--|------------|
| 16. | La Inteligencia Artificial. | 217 |
| | (Ramon López de Mántaras Badía) | |
| 17. | Encontrar vida en otros mundos..... | 223 |
| | (Ester Lázaro Lázaro) | |
| 18. | La gravedad..... | 235 |
| | (Enrique Álvarez) | |
| 19. | El LHC..... | 244 |
| | (Bernardo Adeva Andany) | |
| 20. | Medicina personalizada. | 256 |
| | (Marta Dueñas Porto) | |
| 21. | Computación cuántica. | 266 |
| | (Elías Fernández-Combarro Álvarez) | |
| 22. | Nuclear. | 282 |
| | (Emma López-Alonso Conty) | |
| 23. | Nanotecnología para la salud..... | 296 |
| | (José Miguel García-Martín) | |
| 24. | Sentir el pulso del Universo..... | 309 |
| | (Mariano Santander Navarro) | |
| 25. | El más allá del Modelo Estándar. | 324 |
| | (Xabier Marcano) | |
| 26. | Biomatemática. | 335 |
| | (Antonio Gómez Tato) | |
| 27. | Nueva ruta en la Neurociencia..... | 345 |
| | (Sonia Villapol) | |
| 28. | Exploración espacial. | 353 |
| | (Beatriz Sánchez-Cano y Olivier Witasse) | |

Capítulo 1

La conquista de la Luna: España lo hizo posible.

Es tan grande el desconocimiento de nuestros conciudadanos y de nuestra propia Administración de la vital aportación española a la conquista del espacio exterior, que me veo moralmente obligado a emprender este escrito con el ánimo de añadir luz a esa ignorancia medio secular (50 años).

Al airear estas vivencias profesionales más de cuatro décadas participando en múltiples programas de la NASA, lo hago con el recuerdo vivo en docenas de ambiciosos proyectos espaciales –algunos legendarios- que marcaron un hito en la Historia de la Humanidad.

Pero hoy, en el comienzo del año 2019 en que conmemoraremos el 50º aniversario de la llegada del primer hombre a la Luna, quiero ceñir mis recuerdos al Programa *Apollo*, que de forma tan impactante conmovió al mundo, pero también a los españoles que tuvimos la suerte y el honor de participar directamente en ese ambicioso programa desde las estaciones de seguimiento y control (INTA-NASA) de Fresnedillas de la Oliva y Robledo de Chavela en Madrid, y Maspalomas en Gran Canaria.

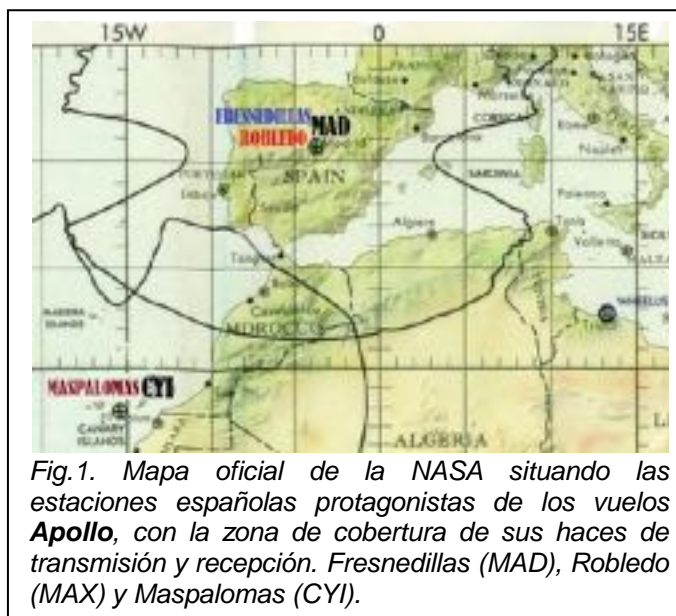


Fig.1. Mapa oficial de la NASA situando las estaciones españolas protagonistas de los vuelos *Apollo*, con la zona de cobertura de sus haces de transmisión y recepción. Fresnedillas (MAD), Robledo (MAX) y Maspalomas (CYI).

Comenzaré por el principio, pues.

Ab initio.

En una de mis repetidas visitas a los archivos del Museo de Historia Americana, en Washington D.C., tuve la sorpresa de encontrarme una copia de un desconcertante documento original del Presidente de los EE.UU., John Fitzgerald Kennedy, que bien se puede considerar el génesis (*ab initio*) de la

Estación de Seguimiento Espacial de Fresnedillas, en la provincia de Madrid. Lo traduzco a continuación porque merece la pena leerlo y entenderlo, e incluyo una reproducción fotográfica del original ilustrando estas líneas.

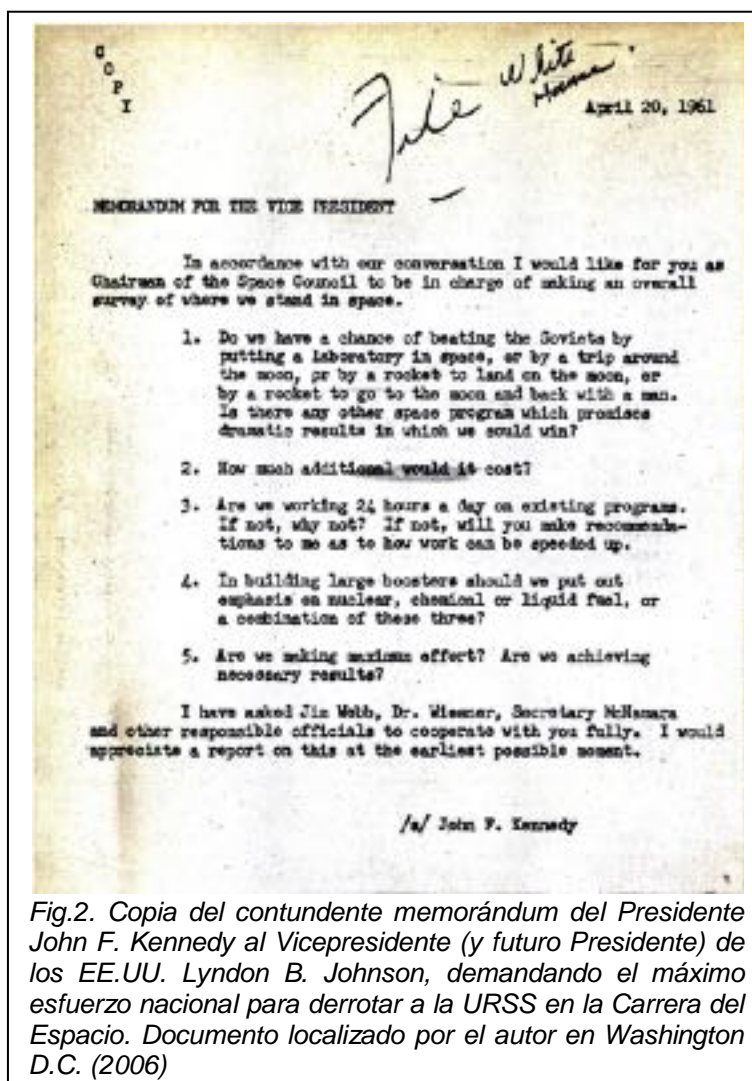


Fig.2. Copia del contundente memorándum del Presidente John F. Kennedy al Vicepresidente (y futuro Presidente) de los EE.UU. Lyndon B. Johnson, demandando el máximo esfuerzo nacional para derrotar a la URSS en la Carrera del Espacio. Documento localizado por el autor en Washington D.C. (2006)

20 Abril, 1961

MEMORANDUM PARA EL VICE PRESIDENTE

De acuerdo con nuestra conversación, me gustaría que Vd., como Presidente del Consejo del Espacio, se encargue de realizar una investigación global sobre nuestra situación en el Espacio.

- 1 *¿Tenemos alguna probabilidad de vencer a los soviéticos, poniendo un laboratorio en el espacio, o haciendo un viaje alrededor de la Luna, o aterrizando un cohete en la Luna, o*

enviando un cohete de ida y vuelta a la Luna con un hombre? ¿Existe algún otro programa espacial que prometa resultados espectaculares y en el que podamos vencer?

- 2 *¿Cuál sería el coste adicional?*
- 3 *¿Trabajamos 24 horas al día en los programas existentes? Si no, ¿por qué no? Si no, ¿quiere Vd. hacerme alguna recomendación sobre cómo se puede acelerar el trabajo?*
- 4 *Para la construcción de grandes propulsores, ¿debemos hacer hincapié en el combustible nuclear, en el químico o en el líquido, o en una combinación de los tres?*
- 5 *¿Estamos haciendo el máximo esfuerzo? ¿Estamos consiguiendo los resultados necesarios?*

He pedido a Jim Webb, al doctor Wiesner, al Secretario McNamara y a otros funcionarios responsables, que colaboren plenamente con usted. Le agradecería un informe sobre esto a la mayor brevedad posible¹.

John F. Kennedy

Cuando leí por primera vez el original en inglés de este escrito, me sorprendió en gran manera el estilo áspero y autoritario del entonces encantador ídolo de las masas John F. Kennedy. Indudablemente, no era ése el talante que transpiraban las imágenes de los noticiarios y telediarios que circulaban de él por el mundo. Qué lejos queda ese *glamour* de los modos en que redactó el memorándum dirigido al Vicepresidente (y futuro Presidente) de los EE.UU., Lyndon B. Johnson.

¹ Nota del autor.

Las tres personas mencionadas por Kennedy en este párrafo eran:

James E. Webb.- Administrador de la NASA desde 1961 a 1968.

Doctor Jerome Wiesner.- Presidente del Comité Consultor de Ciencia de la Presidencia, desde 1961 a 1964.

Robert S. McNamara.- Secretario (Ministro) de Defensa desde 1961 a 1968.

Pero este documento tan especial no estaba sólo, sino arropado por otros miles que conformaban un gran tesoro (al menos para mí), rebotante de información recientemente desvelada a los historiadores al anular oficialmente el Congreso el estatus de “Secreto”, bajo el que habían estado durante tres décadas.

Naturalmente, hice un buen acaparamiento de fotocopias y escáneres de los documentos originales que más llamaron mi atención, elección difícil porque todos ellos daban fe directa o indirectamente de los pasos realizados durante la gestación y desarrollo del ya mítico Programa *Apollo* de la NASA, que corrió paralelo a la mayor parte de mi vida profesional.

Allí encontré también las transcripciones de las conversaciones cápsula-control de tierra y viceversa, que pasaron por Fresnedillas, y que para mí tuvieron un sabor cálido e íntimo por haberlas escuchado en directo mientras la Carrera del Espacio se pespunteaba segundo a segundo. Algunas aparecerán en este escrito más adelante, pero volvamos al guion de esta narración.



Fig.3. Viñeta satírica en la prensa norteamericana del peligroso pulso mantenido entre los líderes de las dos mayores potencias del mundo: Khrushchov y Kennedy a punto de lanzar una bomba “H” sobre el enemigo.

vermo y sin vestigio de vida alguna. Quien vivió aquella época no la olvidará fácilmente.

Fue el 4 de octubre de 1957, cuando los soviéticos pusieron en órbita el primer satélite artificial hecho por el hombre. Se llamó *Sputnik 1* (Acompañante), y causó auténtica conmoción -e incredulidad- en los medios científicos internacionales. Pero apenas un mes después, el 3 de noviembre, el primer ser vivo, la perrita *Laika*, orbitaba la Tierra en el *Sputnik 2*. Pero eso no fue todo, porque también se adelantaron a los norteamericanos con el primer impacto en la Luna con la sonda *Luna 2*, el 12 de septiembre de



*Fig.4. El segundo satélite artificial soviético, **Sputnik 2 (Nov 1957)**, llevó al primer ser vivo al espacio exterior, la perrita **Laika**. Las prisas impidieron diseñar un método de recuperación del cánido. La envenenaron para que no muriera abrasada en la reentrada en la atmósfera.*

1959; y el primer vuelo cercano a Venus con la sonda *Venera 1*, el 12 de febrero de 1961; etc. Pero lo que supuso un auténtico bofetón al orgullo nacional estadounidense, fue la hazaña de Yuri A. Gagarin el 12 de abril de 1961, consiguiendo ser el primer humano en subir al espacio y orbitar nuestro planeta.

La fecha del memorándum para el Vicepresidente, mencionado anteriormente, es una prueba del detonante que debió estallar en la Casa Blanca tras el vuelo de Gagarin ocho días antes. Los asesores del Presidente le orientaron sabiamente, y un mes después, el 25 de mayo de 1961, Kennedy lanzó un órdago en la Sesión Plenaria del Congreso de los EE.UU., en el que expuso con toda firmeza:



Fig.5. El Presidente Kennedy se dirige en sesión conjunta al Congreso y al Senado pidiendo su aprobación para: <<...poner un hombre en la Luna y devolverlo sano y salvo a la Tierra, antes de que concluya esta década (1969).">>

-I believe this nation should commit itself to achieving the goal, before this decade is out, of landing a man on the Moon and returning him safely to Earth. No single space project in this period will be more impressive to mankind, or more important in the long-range exploration of space; and none will be so difficult or expensive to accomplish.-

(Creo que esta Nación debe comprometerse a sí misma a lograr la meta de poner un hombre en la Luna y devolverle sano y salvo a la Tierra, antes de que concluya esta década. Ningún proyecto espacial en este periodo causará tanto impacto en la Humanidad, o será más importante en la exploración espacial a largo plazo; y ninguno será tan difícil o tan caro de lograr.)

Su moción fue aprobada por aclamación, sin fisuras ni titubeos. Si bien la cuestión económica era importante, también lo era el escaso tiempo disponible. La promesa que apoyó el Congreso en Washington, tenía plazo fijo, y había que correr -y mucho-, para poder cumplirla. Mientras tanto, el *premier* Nikita Khrushof había persuadido al Politburó (máximo órgano ejecutivo del Partido Comunista de la URSS) en pleno, de que la potencia que dominara la astronáutica y con ella el espacio, también dominaría el mundo. Asimismo fue respaldado en bloque.

El arriesgado alegato de Kennedy pretendía neutralizar de una vez por todas los fulminantes éxitos soviéticos en la investigación espacial, así que recogió el guante del reto soviético y consiguió la aprobación del presupuesto inicial -y demencial-, de 38.000 millones de dólares, que su antecesor en la Presidencia, Dwight D. Eisenhower había rechazado -asustado-, tan sólo cinco meses atrás (diciembre 1960).



Fig.6. Emblema oficial del Programa Apollo.

Comenzó entonces una frenética galopada en la que el inmenso potencial industrial y técnico norteamericano demostró su imparable ímpetu, pero también se requirió la asistencia científica internacional allá donde la hubiera. Aquella vorágine de entusiasmo y medios económicos sin aparente freno, en la que acabaron interviniendo, directa o indirectamente, alrededor de 400.000 personas de todo el planeta, recibió el nombre de Proyecto *Apollo*. Aquel fue el primer paso para la presencia española en tan ambicioso reto.

Las autopistas electrónicas que comunicaban (y comunican) a los satélites, sondas interplanetarias o naves tripuladas de todo tipo, con los respectivos centros de control en la Tierra eran -y siguen siendo- totalmente esenciales para cualquier vuelo extraterrestre. Cuando se desencadenó la vorágine del Programa *Apollo*, la NASA ya contaba con una red de seguimiento y control para el espacio profundo (*Deep Space Network-DSN*), y otra para satélites y vuelos tripulados orbitales llamada STADAN. En la primera destacaba la de Robledo de Chavela (Madrid), conocida por su código de DSS-61 (*Deep Space Station 61*), y en la segunda la de Maspalomas en la isla de Gran Canaria, imprescindible ésta para decidir la puesta en órbita de los primeros vuelos tripulados norteamericanos de los programas *Mercury* y *Gemini*, o el aborto en su caso. (Desafortunadamente son multitud los compatriotas que ignoran aquella decisiva asistencia española a los primeros vuelos tripulados de la NASA.)



Fig.7. John Glenn visita la Estación de Robledo. (19/10/1965) El astronauta dijo: "No cabe duda de que España ha jugado un importante papel en los vuelos espaciales. Una de las primeras estaciones de seguimiento y observación es la de Robledo de Chavela".



Fig.8. Maspalomas. Antena parabólica de 9m de diámetro. John Glenn, que fue el primer norteamericano en orbitar la Tierra (22/2/1962), reportó desde su cápsula **Friendship 7 (Amistad 7)**: "A la vista, las Islas Canarias ¡qué paisaje!"

Pero ninguna de esas dos instalaciones era plenamente apta ante la complejidad de un vuelo a la Luna, sobre todo si llevaba seres humanos en sus entrañas. La presencia de hombres a bordo de una nave, que por primera vez abandonara el entorno terrestre para hollar otro cuerpo celeste, exigió unas medidas de perfección y eficacia como nunca antes se habían intentado. La NASA anunció que pretendía alcanzar el 100% de seguridad en las múltiples fases del proyecto, y aunque la perfección absoluta no existe, sí se estuvo muy cerca de ella, ya que se consiguió el 99,98%, lo que es poco menos que milagroso si se tienen en

cuenta los miles de situaciones críticas por las que los hombres elegidos habían de pasar.

Llegado a este punto, no puedo omitir la terrible tragedia del *Apollo 1* (27/1/1967), en una simulación en Cabo Cañaveral, en la que murieron de forma horrible los tres miembros de su tripulación, Edward White, Virgil Grissom y Roger Chaffee, delante de sus horrorizados compañeros, que nada pudieron hacer para salvarlos.

Semejante desdicha fue un revulsivo para la NASA, que aumentó las medidas de seguridad, reafirmando entre ellas la creación de una red de seguimiento y control terrestre nueva y específica que, además de cubrir todos los rincones del globo, tuviera al menos tres grandes antenas parabólicas de 26 metros de diámetro, con capacidad para apuntar a cualquier rincón de la bóveda celeste, y procesar debidamente la avalancha de datos que llegarían de la Luna a la Estación, y se transmitirían desde ella, a pesar de la gran distancia Tierra-Luna.



Fig.9. Emblema oficial del que hubiese sido **Apollo 1**, con los nombres de sus tres tripulantes, White, Grissom y Chaffee, muertos trágicamente en una simulación (27/1/1967).

La comunicación directa con los arriesgados viajeros era fundamental para su propia seguridad, así como para garantizar el éxito de la misión encomendada. Sus constantes vitales, su voz e imagen, y las miles de señales que informaban del estado –milisegundo a milisegundo-, de los miles de controles e indicaciones del cohete *Saturno V*, y de las naves en las que harían un periplo de una semana aproximadamente, demandarían una vigilancia exhaustiva por parte del personal especializado en la Tierra. Aquella inestimable ayuda recibió el acertado nombre de “El enlace vital”, y demostró serlo misión tras misión.



Fig.10. El autor a la entrada de la Estación de Fresnedillas, con la antena siguiendo a los astronautas del **Apollo XIV**, que se encontraban en la Luna (también en la imagen) cuando se tomó la foto. (Feb 1971)

La Instalación de Seguimiento de Fresnedillas para Vuelos Espaciales Tripulados (*Madrid Apollo*, para la NASA), nació y se desarrolló en esa dinámica de cuenta atrás, ofreciendo una oportunidad única a los ingenieros, técnicos y especialistas españoles de participar en primera fila en el ambicioso sueño de que un ser humano llegara a pisar otro cuerpo celeste.

La convivencia obligada con los colegas norteamericanos fue personal

e intransferible. Es decir, a cada uno de los alumnos españoles nos fue bien, regular o mal, dependiendo en buena parte de tu propio talante, y desde luego de cómo le cayeras al *teacher* que te habían asignado.

El acuerdo firmado por el Gobierno de España con el de los Estados Unidos, había brindado todo tipo de facilidades para la construcción y manejo –por parte del personal norteamericano-, de las estaciones de seguimiento espacial ubicadas en España. Pero nuestro Gobierno exigió una contrapartida trascendental, la de que el personal técnico español que fuera reclutado, recibiría un adiestramiento intensivo y profundo de los equipos existentes en la estación, así como de los procedimientos a usar en los programas espaciales, y cuando demostrara su perfecta capacitación (*certification*), el *teacher* cedería su plaza al alumno español, y retornaría a su país.



Fig.11. Sello conmemorativo de Correos dedicado a INTA-NASA, iniciativa del autor, con motivo de la aportación española a la Investigación Espacial. Autógrafo del grabador D. Mariano Salamanca, de la Fábrica Nacional de Moneda y Timbre. (1991)



*Fig.12. El comandante de la nave **Apollo VIII**, Frank Borman, visita la Estación de Fresnedillas tras su legendario vuelo alrededor de la Luna, para agradecer el valiosísimo apoyo recibido por su personal técnico. Le atienden Daniel Hunter, Director de la NASA en la Estación, y Jack Zaratzian, Jefe de Operaciones de Bendix. (18/2/1969)*

La realidad del día a día vino a ser más dura que la teoría, ya que los técnicos norteamericanos habían sido contratados por la Bendix, a instancias de la NASA, con unas condiciones económicas envidiables, que de ninguna forma tendrían en los EE.UU. Su renuencia – humana y comprensible-, a perder el envidiable estatus de vida en España, fue un duro escollo adicional que tuvimos que salvar los técnicos españoles de Fresnedillas para obtener la ansiada capacitación. Pero lo conseguimos tras una elaborada “Reconquista”,

(nombre dado por algunos de mis colegas patrios.)

El porqué de la decisión concreta por la NASA de la elección de Fresnedillas de la Oliva para la importante contribución que le esperaba, me demandaría muchas líneas que prefiero emplear en otros eventos. Pero sintetizándolo, señalaré que fue decisivo el hecho de que ya existiera cercana a Fresnedillas la estación de Robledo de Chavela, dotada de una antena de 26 metros de diámetro, que pudiera servir de reserva de la *Apollo Station*. Para ello, se las enlazó con un haz de microondas y por líneas telefónicas, estando ambas alejadas de tendidos de alta tensión, emisoras de radio, líneas de ferrocarril electrificado, o de poblaciones importantes, cuyos servicios pudieran dañar la recepción de las débiles señales que llegarían desde la Luna.



Fig.13. El autor, recién incorporado a NASA-INTA, atendiendo el sistema de Microondas entre las estaciones de Fresnedillas y de Robledo. (1969)

Adoctrinamiento y adiestramiento.

La estación de Fresnedillas fue una gran cátedra para los españoles reclutados por NASA-INTA, porque aprendimos algunas tecnologías y materias que jamás

antes habíamos visto, oído o estudiado en nuestras escuelas y universidades. Desde el primer día de la incorporación de cada uno de nosotros, comenzamos a absorber sin freno lo que el personal norteamericano nos enseñaba (en su lengua, claro), más lo que observábamos en derredor.

Según los conocimientos de cada cual, se nos fueron dando cursos de mayor o menor envergadura, sobre generalidades o sobre temas técnicos concretos. De algunos podíamos saber bastante, de otros menos, y de otros nada en absoluto. Pero a los pocos meses, todos recibimos el *cum laude*, que la NASA llamaba *Certification*, y que te capacitaba para trabajar en nuevas técnicas punta que no paraban de llegar de los EE.UU., aparte de hacerte responsable de uno o más equipos que, indefectiblemente, fueron multiplicándose con el tiempo.

Para los cursos más avanzados, la NASA tenía en el estado de Maryland, un centro de entrenamiento multidisciplinario llamado *Network Test and Training Facility* o NTTF (Centro de Entrenamiento y Pruebas de la Red), donde se impartían cursos especializados de disciplinas tales como: informática, ordenadores, radiofrecuencia, microondas, servomecanismos, hidráulica, criogenia, comunicaciones, y así un largo etcétera, que tenían aplicación directa en los equipos operativos que la NASA había depositado en nuestras manos.

Tan pronto como un nuevo técnico español era captado para Fresnedillas, se le asignaba el equipo o sección más acorde con sus conocimientos, se le ponía al día sobre el vastísimo Programa *Apollo* (*Apollo Indoctrination*), y tan pronto era posible, se le enviaba al complejo educacional NTTF. Allí recibiría un denso curso de uno a tres meses de duración, sobre uno o varios equipos, que a su vuelta a Fresnedillas estarían bajo su tutela y responsabilidad.



Fig.14. Fresnedillas. En primer plano consola de Receptores en la sección de Radiofrecuencia, y al fondo la consola del Servo (Antena). Todo el equipo en funciones es español. (1971)

Rara era la semana en que no hiciéramos simulacros y pruebas internas, que luego se ampliaban al resto de la red de estaciones y centros de control de la NASA en todo el planeta. Así, cuando llegaba un lanzamiento de verdad, todo transcurría como si fuera un simulacro más, es decir, perfectamente.

“Madrid is Green and Go!”, era nuestra respuesta cuando inquirían sobre nuestro estatus los centros de control de Goddard y

Houston. Era mucho más concisa que su traducción al español, que sería: “Madrid – Fresnedillas tiene todos los equipos funcionando y su personal está dispuesto para todo”.

Detrás de la exclamación “*Green and Go!*”, se ocultaban las agotadoras pruebas que desarrollábamos entre *Apollo* y *Apollo*, calibrando equipos, disponiendo configuraciones, enviando y recibiendo señales, y ¡cómo no!, preparándonos a nosotros mismos para el siguiente vuelo. Repasábamos una y otra vez los voluminosos manuales que la NASA actualizaba con gran frecuencia, introduciendo cambios que, a veces desbarataban todo lo que uno creía ya más que sabido.

Recuerdo muy bien que en una ocasión se detuvo la cuenta atrás de un lanzamiento a la Luna (*Apollo XVI*), porque las dos estaciones de Madrid dejaron de estar disponibles a tan sólo 9 días del despegue hacia nuestro satélite. Ocurrió el viernes 7 de abril de 1972 cuando se cortaron totalmente las comunicaciones por microondas entre Fresnedillas (*Madrid Prime*) y Robledo (*Madrid Wing*).

Fuertes rachas de viento surgidas inesperadamente, parecían ser las causantes del desafuero. El anemómetro instalado en el tejado del edificio de Operaciones indicaba velocidades de más de 160 km/h, y si esa lectura se tomaba a 4 metros de altura, ¿qué no ocurriría allá arriba en *La Almenara*, a 1.259 metros de altura, donde se encontraba el repetidor de microondas que enlazaba a las dos estaciones mencionadas?

Una vez informada la NASA, su decisión fue rotunda: si las dos estaciones *Madrid Apollo* no podían funcionar a pleno rendimiento, se detendría la cuenta atrás del lanzamiento del *Apollo XVI*. ¡Así de valiosa era nuestra función!

Inmediatamente se organizó un grupo de técnicos (entre los que se encontraba este autor) para escalar aquella cima, cargados de herramientas y equipos como los *sherpas* tibetanos, para subsanar el posible daño causado por el inesperado huracán a la torreta que sostenía dos platos parabólicos de 1,5m de diámetro cada uno.

No voy a detallar la dureza de la ascensión, ni los riesgos de caer despeñados, porque lo importante ahora es constatar que logramos nuestro propósito, aunque nos dejáramos amplias zonas de nuestra epidermis en ello. Nada más descender y retornar a Fresnedillas, informamos a nuestros superiores, que satisfechos, trasladaron a Goddard y a Houston la buena nueva de que Fresnedillas (*Madrid Prime*) y Robledo (*Madrid Wing*), volvían a estar operacionales al cien por cien. ¡El *Apollo XVI* podía reiniciar su cuenta atrás! Despegó sin más problemas el 16 de abril de 1972, y fue un gran éxito, como es ampliamente sabido.



Fig.15. El astronauta Charles Duke atiende las explicaciones del autor sobre el Museo Lunar de Fresnedillas. Duke fue el **CapCom** del **Apollo XI** y piloto del módulo lunar del **Apollo XVI**.



Fig. 16. Avión Lockheed C-121 **Constellation**, uno de los cuatro que modificó la NASA simulando ser una nave *Apollo*. Sobrevoló varias veces la nueva Estación de Seguimiento Espacial de Fresnedillas, para su puesta a punto en el **Programa Apollo**.

Pero volviendo a los repetidos entrenamientos, la NASA, para que todo fuera lo más real posible, se había hecho con una flotilla de cuatro aviones Lockheed C-121 *Constellation*, que fueron transformados interiormente en auténticas cápsulas espaciales, con sus ordenadores, transmisores y receptores, equipos de navegación, telemetría, moduladores de voz, etc. El equipo personal se componía

de doce técnicos (uno de ellos hacía las veces de astronauta), más tres tripulantes (dos pilotos y un navegante).

El *Constellation* comprobaba exhaustivamente la eficacia de nuestro personal sobrevolando repetidas veces las estaciones de Fresnedillas y Robledo, lo que provocó cierto nerviosismo en algunos pueblos de la sierra de Guadarrama, por confundirlo sus ciudadanos con un OVNI. La prensa se hizo eco de esos “inexplicables” avistamientos, mientras nosotros sacábamos el máximo provecho a aquella nave *Apollo* de pega.

Traqueando.

Hasta yo mismo me horrorizo de la palabreja que da título a este subcapítulo: *Traqueando*. Era una de las más usadas de nuestro cotidiano argot laboral, fruto de una fusión morfosintáctica y semántica del español con el inglés, conocida como *espanglish* o *ingañol*, a elegir. La realidad es que nos permitíamos todo tipo de atrocidades lingüísticas durante la faena, siempre que no nos oyeran nuestros colegas norteamericanos de NASA-Bendix.



Fig. 17. La consola del Servo (Antena) de Fresnedillas en plena vorágine del vuelo **Apollo XI** a la Luna. Cuantos más ojos y más oídos mejor. En primer plano el norteamericano Lenny Parker con dos compatriotas. (Julio 1969)

Cuando en Fresnedillas decíamos que *traqueábamos*, queríamos señalar que: “buscábamos” o “rastreamos” por el espacio una nave, sonda, o satélite, la “localizábamos”, “capturábamos” su señal, continuábamos

“persiguiéndola”, allá por donde navegara, mientras le “sacábamos” toda la información que había conseguido por aquellos cielos de Dios, a la vez que le “dábamos instrucciones” (*commands*) para que se comportara debidamente durante la siguiente órbita u horas de singladura espacial. Eso era el *Tracking*, ni más ni menos. De ahí que involuntaria y cómodamente optáramos por el término aberrante de *traquear*, que englobaba la actividad polivalente mencionada.

Recordaré al lector no avezado, que el ciclópeo *Saturn V*, tras despegar de Cabo Cañaveral, soltaba sus dos primeras fases en breves minutos, una vez consumido su combustible, y el resto del cohete portando las naves *Columbia* e *Eagle*, entraba en órbita de aparcamiento terrestre -que así se llamaba-, para comprobar minuciosamente el funcionamiento de todos los equipos de a bordo, con las instrucciones que los astronautas constataban con el personal de tierra. Aquellas eran horas de auténtico fragor de batalla, donde nada podía quedar al albur, y éramos conscientes de que cualquier fallo de equipo o error del personal técnico, podía dar al traste con la misión, o incluso poner en peligro la vida de los astronautas.



Fig.18. Diagrama del gigantesco cohete **Saturno V**, obra maestra del ingeniero de origen alemán **Wherner von Braun**. De 110 metros de altura, fue el mayor jamás construido. Llevó a 27 hombres hasta la Luna sin un solo fallo.



Fig.19. El Centro de Control de la Misión (MSC), en Houston en plena efervescencia durante el vuelo del **Apollo XI**. La coordinación de todos los ingenieros y especialistas, junto a la conexión continua con las estaciones de seguimiento en todo el globo, hicieron posible el éxito histórico del Programa **Apollo**. (1969)

Cuando por fin el *Mission Control Center* (Centro de Control de la Misión) en Houston (Texas), con la aquiescencia de la tripulación del *Apollo*, decía las palabras mágicas *Gofor TLI!* (¡Adelante con la Inserción Trans Lunar!), un último bramido del cohete *S-IV-B* vencía la gravedad terrestre y ponía rumbo a nuestro satélite, al que llegaría unos tres días después. A partir de ese momento, en Fresnedillas el frenesí iba amainando y nosotros permitíamos que nuestros riñones se apoyaran

suavemente en el respaldo del asiento, donde habíamos permanecido envarados como estacas durante las últimas horas.

El sueño de una noche de verano.

Una de las preguntas más repetidas que me han hecho -y me siguen haciendo- los asistentes a mis conferencias, es para saber qué emociones experimentamos mis compañeros y yo en Fresnedillas, en los momentos históricos de la llegada del primer hombre a la Luna. La respuesta suele defraudar porque ninguno de quienes participamos en aquellos hechos tuvimos tiempo para hacer disquisiciones filosóficas sobre el alcance de la evidente hazaña conseguida en nombre de la Humanidad.

La noche del 20 al 21 de julio de 1969, el equipo técnico de Fresnedillas repetía una vez más lo que venía haciendo desde meses atrás, siguiendo las mismas pautas, ejecutando los mismos pasos según lo



Fig.20. Sello conmemorativo del US Post Office, emitido por la llegada del primer hombre a la Luna. Se hicieron 152 millones de ejemplares, que no llegaron a cubrir las peticiones de los coleccionistas de los cinco continentes. (1969)

previsto y estudiado cientos de veces. El personal sólo tenía ojos y oídos para escrutar ávidamente las indicaciones que mostraban sus equipos, comprobar que los valores aritméticos o eran exactos, o estaban dentro de un margen aceptable. Cada técnico formaba un indisoluble tándem con su equipo, como si fueran dos entes en uno solo, y ése era su único mundo, que no era poco.

El momento para reflexionar sobre el conjunto de la labor felizmente desarrollada en Fresnedillas, llegaba sólo cuando el relevo venía a levantarte de la silla para continuar tu labor, y tú permanecías aún unos minutos de pie detrás de él tanteando tu bagaje de adrenalina y asegurándote de que todo proseguía como la seda. Eran la prensa y los informativos de televisión quienes divulgaban como una notoria gesta la aportación de tu saber, eficacia y temple de nervios.

Salvando las comparaciones, y desde luego la enorme distancia Tierra-Luna, me permito reproducir aquí el comentario que le oí a Edwin (Buzz) Aldrin en 1989, en el 20º aniversario de su llegada a nuestro satélite, cuando le hicieron una pregunta similar: <<What did you feel when you first stepped on the Moon? What did you think on such historical moment? >> (¿Qué sintió cuando piso la Luna por primera vez? ¿Qué pensó en ese momento histórico?)



Fig.21. Los héroes del Apollo XI. (Izda. a dcha.): Neil A. Armstrong, Michael Collins y Edwin "Buzz" Aldrin. (1969)

Ruego al lector que me disculpe por poner el texto en inglés, pero ya que tengo en mi poder una transcripción literal de sus palabras en aquella teleconferencia, me debo al rigor de sus expresiones originales. Aldrin se expresó así:

<<There wasn't time to savour the moment. It seemed as though what we were doing was so significant that to pause for a moment and reflect metaphysically was really contrary to our mission. We weren't trained to smell the roses. We weren't hired to utter philosophical truisms on the spur of the moment. We had a job to do. >>

(No había tiempo para saborear el momento. Nos parecía que lo que estábamos haciendo era tan importante, que parar un momento para reflexionar metafísicamente iba en contra de nuestra misión. No fuimos entrenados para oler las rosas. No se nos había contratado para proferir perogrulladas en aquel momento concreto. Teníamos una tarea que cumplir.)

Es evidente que Aldrin y yo pensábamos igual.

El diario ABC había dicho unos días antes, el 10 de junio de 1969, en **negrita**:

Plena confianza de la N.A.S.A. en los técnicos españoles. (Seguía más adelante):

<<... reconocimiento oficial al hecho de que el personal que trabaja en ella, directivos, técnicos, administrativos y subalternos, todos ellos empleados del I.N.T.A., tienen el alto nivel técnico necesario para la realización de su misión. >>



Fig.22. El CapCom del **Apollo XI** Charles Duke en acción en el Centro de Control de la Misión (MSC), en Houston. Tras él los astronautas del **Apollo XIII**, Jim Lovell y Jack Swiggert. (Julio 1969)

Sería falsa modestia contradecir al **ABC**, porque el devenir de aquel vuelo lunar le dio la razón. Todo salió perfecto en las estaciones de seguimiento españolas. ¿Además quiénes éramos nosotros para desdecir a tan ilustre y decimonónico periódico?

Los tripulantes del *Apollo XI*, mejor dicho, del módulo lunar *Eagle* (Águila), no descollaban por parlanchines precisamente. Mantenían largos silencios que

quebraba de vez en cuando el *CapCom* (astronauta Charles Duke) desde Houston, simplemente para oír su voz. Nosotros en Fresnedillas no le quitábamos ojo a las registradoras de papel (*strip chart recorders*), donde veíamos en directo la respuesta de sus electrocardiogramas, encefalogramas, presión sanguínea y demás constantes vitales. El ritmo cardiaco era normal y su actividad cerebral la esperada en cada momento. Aun así se echaba en falta la cháchara típica de otros astronautas que mantenían la línea de voz con la Tierra calentita, aunque fuese con frases monótonas y aburridas. El caso era hablar y ser contestado. La recepción de la voz humana en la soledad cósmica tenía para los astronautas propiedades mágicas como el bálsamo de Fierabrás para Don Quijote.



Fig.23. Uno de los frecuentes electrocardiogramas de los astronautas del **Apollo XI** recibidos en Fresnedillas. Aldrin (72 ppm), Armstrong (108 ppm), y Collins (84 ppm). (Julio 1969)

Pero ni Armstrong ni Aldrin estaban por la labor de amenizar el viaje, y el aburrido Collins, allá arriba fue quien quebró el silencio muchas veces con chanzas y chirigotas, quizás porque en su soledad necesitaba más que nadie el calor humano, aunque fuera exclusivamente verbal. Hasta el diario vespertino *Pueblo* había editado una crónica el 19 de julio, titulada *Tripulación aburrída*, precisamente por los cansinos silencios espaciales.

Por la posición de la Tierra en el momento cumbre del alunizaje, le correspondió a Fresnedillas ser los oídos y la voz del centro de control de Houston. Ese mismo día épico del 20 de julio, el director de vuelo Eugene F. Kranz, ordenó despertar a la tripulación del *Apollo XI* con una selección de briosas marchas de John Philip Sousa, tan conocidas de todos, por ser las habituales de los encuentros deportivos y actuaciones circenses. Kranz quería a la tripulación marchosa para el gran acontecimiento.

Luego siguieron las fases de frenado para aliviar el tirón de la Luna y quedar atrapados en su órbita, más tarde la separación de los módulos de mando y servicio (CSM) del módulo lunar (LM), alejando a Armstrong y Aldrin de su colega Collins, que pasó a ser el hombre *más sólo del universo desde Adán*, como se autodenominó unas horas después.

Cuando el módulo lunar *Eagle* descendía a 1.800 metros hacia su destino en la superficie lunar, el ordenador de a bordo encendió una alarma de color ámbar, la 1202. Todo el mundo se echó a temblar, porque en esas circunstancias trepidantes cualquier anomalía podía suponer un desastre. Unos minutos después volvieron a aquietarse los corazones al asegurar un informático empollón del control de Houston, que la 1202 era un simple aviso del exceso de actividad del ordenador. El pobrecito ordenador estaba desbordado y simplemente se quejó. El *CapCom* y futuro astronauta del *Apollo XVI* Charles Duke, tranquilizó a los chicos de arriba diciéndoles que siguieran adelante, que todo estaba bajo control. En Fresnedillas también se exhaló un suspiro, pero por poco tiempo, porque apenas doce minutos más tarde el módulo *Eagle* sobrevolaba la zona prevista para el aterrizaje.

Aldrin iba cantando una retahíla de números (como los niños de San Ildefonso en Navidad, pero sin alegría), haciéndonos saber dónde se encontraba la nave cada pocos segundos. Aunque Aldrin lo decía en inglés y en su sistema de medidas, me he permitido hacerlo asequible para quienes vivimos en el sistema métrico decimal:

<<...250 metros [altitud], reduciendo a 25 [kilómetros por hora],...200 metros, reduciendo a 21 [km/h]...165 metros, reduciendo a 16 [km/h]...121 metros, reduciendo a 10 [km/h], hacia adelante 100 metros, descendiendo a 4 km/h...>>

De repente los navegadores de Houston percibieron que algo iba mal, el módulo lunar había dejado prácticamente de descender pero de repente entre 60 y 100 metros de altitud aumentó su velocidad a ¡unos 88 km/h! ¿Dónde iban

esos locos? El *CapCom* Charles Duke avisó: <<*¡60 segundos!*>>. Ese era el tiempo de consumo de combustible que les quedaba para abortar la misión y encender el motor que les traería de vuelta a la Tierra, si no conseguían posarse suavemente en suelo llano y en posición vertical. Su retorno dependía totalmente de cómo aterrizaran. Ese fue otro de los momentos en que todos, en Houston y en Fresnedillas contuvimos de nuevo el aliento.

La retahíla de Aldrin se reanudó:

<<...bajando despacio...61 metros, reduciendo a 1,5 [km/h]...5% de combustible [queda en la reserva],...23 metros, encendemos luces, ...reduciendo a 0,8 [km/h] , ...12 [metros], levantamos polvo, ...9 [metros], sombra tenue, 1,3 [km/h] avante, un poco a la derecha,...va bien...>>

Por el transmisor de Fresnedillas salió el aviso del *CapCom* que nos inquietó a todos: <<*¡30 segundos!*>>, alertándonos de que no buscaran más el aparcamiento perfecto, porque el combustible estaba en las últimas. Pero los segundos siguieron pasando en un silencio agobiante que nadie en la Tierra se atrevía a interrumpir. Todos nos mirábamos sin decir palabra y sin quitar ojo a los relojes que parecían correr desenfrenadamente. Detrás del *CapCom*, el Director de la Misión *Apollo XI*, George Hage, rezaba en voz baja: <<*¡Bájalo, Neil! ¡Bájalo!*>>

Por fin se encendieron las luces indicadoras de que las cuatro patas del módulo lunar habían tocado suelo, y la voz de Aldrin, vino a relajarnos a todos: <<*Contact Light...okay, engine stop!*>> (¡Luz de contacto. OK. Motor parado!), cuando sólo les quedaban 10 segundos más de propergol. ¡Uf! Estuvieron muy cerca de la catástrofe, pero el baqueteado Armstrong, al no localizar el lugar previsto estudiado en la Tierra, se hizo con los mandos manuales y buscó sobre la marcha el lugar menos escabroso posible, y allí decidió posar la nave voladora más antiestética que se haya conocido jamás.

De nuevo la voz del comandante Neil Armstrong cruzó el espacio, y cambiando su tono aséptico por otro algo más vivo, nos paralizó a todos:

<<Houston, Tranquility Base here. The Eagle has landed.>>

(Houston, aquí Base de la Tranquilidad. ¡El Águila ha aterrizado!)

El griterío exultante de la conocida sala de control de Houston se oyó en Fresnedillas por las líneas de comunicación, fundiéndose con las de nuestros técnicos americanos y españoles, que desbordamos la emoción contenida durante los últimos densísimos minutos. El *CapCom* Duke pudo por fin hilar

palabra y contestar a los nuevos selenitas lo primero que se le ocurrió, que era el fiel retrato de lo que acababa de ver a su alrededor:

<<Roger, *Tranquility*. We copy you on the ground. You've got a bunch of guys about to turn blue. We're breathing again. Thanks a lot. >>

(Enterado, *Tranquilidad*. Os oímos desde tierra. Había aquí a un montón de tíos a punto de ponerse azules. Ahora volvemos a respirar. Muchas gracias.)



Fig.24. Sello del módulo lunar *Eagle* del *Apollo XI*, emitido por el Emirato de Manama.

Eran las 9 y 18 minutos de la noche del sábado 20 de julio en la España peninsular.

Alguien nos comentó que la emoción también se había desbordado en la Sala de Prensa, donde los enviados especiales entraban y salían presurosos para comunicar la buena nueva a sus editoriales, familias y amigos. No era para menos.

En esos agitados minutos, alguien en los EE.UU. visitó el cementerio de Arlington en Virginia, y depositó un ramo de flores en la tumba del presidente Kennedy, verdadero promotor y motor del Programa *Apollo* hasta que le asesinaron en 1963. Las flores iban acompañadas de una nota que decía: *-Mr. President, the Eagle*

has Landed.- (Sr. Presidente, el *Águila* ha alunizado.) Seguro que John F. Kennedy lo agradeció.

Unos mil millones de televidentes de los cinco continentes, esperaban que en pocos minutos aparecieran los astronautas en escena hollando la virginal Selene, pero no, hubo que esperar varias horas. Como estaba escrito, Armstrong y Aldrin se pusieron inmediatamente a preparar el módulo lunar por si tenían que volver a toda prisa ante cualquier imprevisto no deseado. No en balde, estaban en otro mundo y nunca se sabe... Emplearon en ello casi dos horas, y con los nervios a flor de piel manifestaron a Houston su deseo de salir al exterior sin más dilación, pero los directores, con la cabeza más fría y mucho más responsables que los astronautas en ese momento, aceptaron a regañadientes saltarse la agenda prevista y permitirles salir al exterior **tres horas más tarde**, que tiempo tendrían de dormir en su viaje de vuelta a casa.

Pero antes de salir al exterior, el piloto del *Eagle* nos sorprendió a todos pidiendo al control de tierra un par de minutos de silencio y meditación. Nos miramos unos a otros con cara de interrogación, intentando adivinar el sentido

de las palabras de Buzz Aldrin, que fueron exactamente estas, que transcribo del documento correspondiente:

"This is the LM pilot. I'd like to take this opportunity to ask every person listening in, whoever and wherever they may be, to pause for a moment and contemplate the events of the past few hours and to give thanks in his or her own way."

(Soy el Piloto del Módulo Lunar. Me gustaría aprovechar esta oportunidad para pedir a todos los que me están escuchando, sean quienes sean y estén donde estén, que hagan una pausa momentánea y repasen los acontecimientos de las últimas horas y den gracias por ello a su modo.)



*Fig.25. Bolsita privada de Aldrin en la que llevó el cáliz, el vino y un trocito de pan bendito para comulgar en la Luna. Lo hizo ya posado el **Águila** en el Mar de la Tranquilidad, antes de salir al exterior.*

Después de su retorno a la Tierra, supimos por el mismo protagonista del inesperado acto sacramental, que tras pedir al comandante Neil Armstrong la preceptiva autorización, comulgó con el cáliz y pan bendecido que había llevado consigo secretamente en un bolsillo del brazo derecho. La NASA no vio con buenos ojos aquel acto, ya que el año anterior la tripulación del *Apollo VIII* había leído en una retransmisión televisiva unos párrafos del Génesis bíblico, acto que fue llevado a juicio tras una denuncia de un grupo religioso radical estadounidense.

El mensaje pseudo crítico de Aldrin respondía al requerimiento de la NASA de que sus astronautas no hicieran manifestaciones ni actos religiosos que pudieran provocar críticas de cualquier orden. De ahí que en Fresnedillas no entendiéramos a priori el mensaje de Aldrin.

El sufrido Collins fue el patito feo de aquella proeza, porque no sólo estuvo a las puertas de la gloria, sino que la acarició una y otra vez mientras orbitaba nuestro satélite, oyendo cómo sus dos compañeros allá abajo se apropiaban de toda la fama del momento. El equipo humano del *Apollo Wing* en Robledo, dedicó toda su atención en exclusiva al casi olvidado piloto del módulo de mando *Columbia*, Michael Collins. En sus órbitas lunares, Collins sólo disponía de 7 fugaces minutos para charlar directamente con sus añorados compañeros Armstrong y Aldrin, en los que apenas tenían tiempo para cruzarse un saludo en plan telegráfico.

El implacable movimiento de rotación de la Tierra hizo que la Luna se ocultara por el horizonte de poniente de Fresnedillas (*Apollo Prime Site*) y de Robledo (*Apollo Wing Site*), perdiendo la primera la señal del *Eagle* con Armstrong y Aldrin dentro, y la segunda la del *Columbia*, pilotado por Collins.

Cuando el Control de la Misión en Houston nos relevó a ambas estaciones de nuestra responsabilidad, empezamos a notar un evidente cansancio, contenido hasta ese momento, pero ninguno quisimos volver a nuestras casas. ¿Quién se podía perder el mayor acontecimiento de nuestra vida? ¿Quién iba a dejar que el sueño nos impidiera ver en directo el primer paso de un ser humano en otro cuerpo celeste? Así nos dedicamos a comentar detalles de las últimas horas, haciendo corrillos o deambulando por las salas de Operaciones sin rumbo fijo, saboreando el momento, aunque atentos a las conversaciones entre Houston y el trío de la fama para asegurarnos de que todo seguía según el plan previsto (*schedule*, en inglés).

Me acerqué a la sala de Prensa, ubicada en el edificio de Cafetería, y la encontré rebosante de corresponsales, VIPS y visitantes autorizados, que observaban los monitores de televisión que mostraban las imágenes que Prado del Rey transmitía a todos los españoles. Otro recibía la información de la oficina de Relaciones Públicas de la NASA en Houston, y otros dos mostraban aspectos de las salas de Operaciones, ahora ya dispensadas de obligaciones.

Al batiburrillo de periodistas, se sumó aquella noche en Fresnedillas el embajador norteamericano Mr. Robert C. Hill, acompañado de una aristocrática belleza muy en boga entonces, Ira von Fürstenberg, Princesa de Hohenlohe.

Después de que el Sr. Embajador vio lo que vio y oyó lo que oyó en Fresnedillas, le faltó tiempo para manifestar su satisfacción a un periodista de la agencia Cifra, que publicó lo siguiente al día siguiente del feliz aterrizaje del *Apollo XI*:

- Acaba de empezar una gran empresa humana. Hemos podido presenciar todos el comienzo de una aventura audaz, pero no temeraria. Pero no vuelan solos. Vuelan apoyados por los equipos de técnicos que siguen su trayectoria segundo a segundo desde la Tierra, muy singularmente desde la estación espacial en Fresnedillas, casi a las puertas de Madrid, pues le acompañan los buenos deseos, la admiración y las orientaciones de la mayor parte de la raza humana, que se sabe representada por estos tres hombres.

Las funciones de la estación espacial del “Apolo” en Madrid, que abarca la tercera parte del cielo, son de la máxima importancia para este vuelo y para todo el programa “Apolo”.

Deseo expresar mi sincero agradecimiento al Gobierno español – termina diciendo el embajador norteamericano-, por su muy apreciada colaboración y a los competentes técnicos espaciales españoles por su intenso trabajo y consagración al éxito de esta empresa espacial.- Cifra.-

¡Un hurra por el señor Embajador!

Llegó el momento tan deseado. La NASA, a través de la red de seguimiento MSFN, nos ofreció imágenes de la cámara de televisión sujeta al fuselaje del módulo lunar. Era un encuadre estático de la escalerilla del módulo lunar por donde había de descender Armstrong. La imagen era en blanco y negro y escasamente nítida, pero más que apta para lo que un tercio de la Humanidad estaba a punto de contemplar.



Fig.26. Dibujo a carboncillo de Paul Calle, autor del sello conmemorativo de 10 centavos, según se imaginó cómo Neil Armstrong hollaría la superficie lunar

De pronto una mancha deforme, correspondiente al aparatoso atuendo de supervivencia de Armstrong empezó a cruzar la pantalla lentamente de arriba a abajo. Parecía un fantasma deforme con su aura grisácea, que temiera despeñarse por la escalerilla abajo. El pie izquierdo, embutido en una enorme bota, buscaba con cuidado el siguiente escalón, y así hasta que llegó al último, desde donde se dejó caer a la cazoleta que rebordeaba la pata del *Eagle*.

A las 109:24:48 de vuelo del *Apollo XI*, es decir, las 03:56:15 de la madrugada (hora española) del domingo 21 de julio, los españoles aún despiertos pudieron ver como en un lejano sueño a Armstrong poner al fin los dos pies en el suelo lunar, convirtiéndose en el primer terrestre en invadir otro cuerpo celeste.

El audio de Neil Armstrong también dejaba que desear, pero aun así pudimos intuir, más que oír sus palabras, injertado en aquel muñeco hinchado y torpe sacado de un anuncio de Michelin, que tantísimas veces hemos oído y repetido después:

<<*That's one small step for (a) man; one giant leap for mankind.* >>

(Éste es un pequeño paso para un hombre; una zancada gigante para la Humanidad.)

En la sala de Comunicaciones de Fresnedillas, uno de los teletipos comenzó a hacerse oír con su escandaloso traqueteo adornado con la ráfaga de campanilleos que marcaban el final del mensaje: ¡Clinc!, ¡ Clinc!, ¡ Clinc!, alertando al operador cercano de que acababa de recibirse un nuevo mensaje. La escritura sobre papel con seis copias se sucedía sin freno. Los textos procedían de las agencias de noticias más famosas pretendiendo ganar la primicia que ya conocíamos cientos de millones de personas en todo el globo. En todos los idiomas posibles y con redacción telegráfica anunciaban a la rosa de los vientos el gran acontecimiento: el hombre había hollado la Luna. Fue la crónica más breve, y a la vez la más grande jamás dada por aquellas infernales máquinas, sustituidas ahora por la también demoniaca Internet.

Se decía que el periodismo era el arte de contar las cosas con el mínimo de palabras posibles. Algunas de las más importantes agencias informativas dieron así de breve la noticia, con apenas cuatro palabras la *Associated Press* y la *Reuter*, y siete palabras –la más larga- la agencia *France Press*:

up 77 c
accapollo 7/21
bulletin
~~bulletin~~
armstrong sets foot on moon
upi spaceflight
sp0358
-.-.-.-.-.

fl a s h
el hombre pisó la Luna
de Nueva York. 03.56
-.-.-.-.-.

a345ho
zyyuivwyyf
flash
space center
armstrong steps onto moon
tnpap

afp 079
f l a s h Ω Ω Ω Ω Ω Ω Ω Ω
L 'homme pose le pied sur La Lune.
afp. 03.56

Antes de salir al exterior, Aldrin había reclamado a Armstrong un trueque. Ya que iba a ser Armstrong (contra el pronóstico inicial) quien pasara a la inmortalidad por ser el primer hombre en pisar otro cuerpo celeste, Aldrin quería ser el protagonista casi absoluto de las fotografías que se obtuvieran

durante su permanencia en nuestro satélite, y para ello le entregó la cámara Hasselblad para que se la colgara al cuello y la utilizara sin contención.

Dieciocho minutos después del famoso *pequeño paso para un hombre*, Aldrin asomó por la escotilla e inició el descenso, teniendo antes la precaución de dejarla entornada y obstruida, no fuera a cerrarse tras ellos, y tuvieran que quedarse en la Luna para siempre, ya que a los ingenieros diseñadores de la empresa Grumman no se les había ocurrido dotar a la escotilla de un manillar exterior. La feliz idea fue aplaudida por Armstrong, como pudimos escuchar en las líneas de *Air-to-Ground Voice*:

<<Aldrin: *Now I want to back up and partially close the hatch, making sure not to lock it on my way out.*-Armstrong: *A particularly good thought.*>>

(Aldrin: Ahora quiero volver a subir y entornar la escotilla, asegurándome de que no la cierro mientras salgo.- Armstrong: Una idea especialmente acertada.)

Desde Houston, con todo el tacto posible, insistían una y otra vez en que los dos turistas torpones se pusieran a recoger piedras sin dilación, en vez de extasiarse ante aquella <<*magnificent desolation!*>> (¡grandiosa desolación!), como la definió Aldrin en su primer vistazo.

Tantearon su extraña movilidad y liviandad dando saltitos como canguros, y Armstrong comenzó a buscar rocas en medio de aquel extenso colchón de *polvo de talco* pardusco. Aunque llevaba una bolsa de teflón para guardarlas, sus dedos, agrandados y deformes por los guantes, eran incapaces de abrirla y optó por irlas metiendo en los bolsillos de las perneras del pantalón.

El Presidente Richard Nixon quiso tener también su porcentaje de protagonismo en aquella aventura, y les llamó por teléfono, recalcando que lo hacía desde el famoso Despacho Oval de la Casa Blanca.

Apremiados por Houston, Armstrong y Aldrin se replegaron al *Eagle*, tras pasar apenas 2 horas y 47 minutos pululando por la Luna. Desplegaron la bandera de las barras y estrellas, los equipos del EASEP (Early Apollo Scientific Experiment Package – Primer conjunto de experimentos científicos del Apollo), que dejaron funcionando, y cuyas señales se recibieron en la Tierra inmediatamente. Y dentro de una bolsa, iba un disco de silicio con mensajes grabados de 73 jefes de Estado (con una incomprensible ausencia, la del Jefe del Estado español); un simbólico broche de oro con forma de ramita de olivo; un emblema bordado del *Apollo 1* con los nombres de sus tres astronautas trágicamente fallecidos: White, Grissom y Chaffee.

Pero el detalle más emocionante fue cuando depositaron con cuidado ritual en el suelo lunar, dos medallas que les habían hecho llegar las viudas de los cosmonautas soviéticos Vladimir Komarov y Yuri Gagarin, también desaparecidos trágicamente.

Sus esposas habían pedido que quedaran en la Luna para siempre, y allí están. Allí quedó también una placa metálica abrazada a una de las patas del módulo lunar *Eagle*, que mostrará eternamente un escueto mensaje firmado por Armstrong, Aldrin, Collins y el Presidente Nixon, que difícilmente hubiese podido mejorarse:



Fig.27. Placa metálica fijada a una de las patas del módulo lunar *Eagle* del **Apollo XI**, con un mensaje de buena voluntad con las firmas de los tres astronautas, más la del Presidente Nixon.

<<Aquí, hombres del planeta Tierra pusieron por primera vez el pie en la Luna. Julio 1969. (Era Cristiana) Vinimos en paz en nombre de toda la Humanidad. >>

Cuando los dos turistas astrales recibieron la orden de *<<gofor relax>>* (acomodaos para dormir) en el *Eagle*, pensé que era el momento de imitarles, tras tantas horas de estresante vigilia. Me entretuve en el pasillo para ver en los monitores allí instalados, cómo reponían las escenas de los primeros pisotones de Armstrong. Se arrimó a mí un *janitor* (auxiliar de la limpieza) ya mayor, oriundo de uno de los pueblos serranos cercanos, y apoyándose con cierta apostura en su mocho me preguntó, a la vez que indicaba con un gesto de cabeza las imágenes grises que había enviado el *Apollo XI*: *<< ¿Usted de verdad se cree que hay hombres en la Luna? Pues yo no. >>* Y ante la cara de asombro que le debí poner, y sin darme opción a responder algo, añadió: *<< No caben. >>* Anécdota inolvidable, que por eso la traigo aquí.

Y siguiendo el hilo de las incredulidades llamadas por algunos “teorías conspiratorias” insistiendo en que los norteamericanos jamás llegaron a la Luna, me permito dos breves digresiones: 1) si aquello hubiese sido una farsa –que no lo fue-, los soviéticos lo hubiesen sabido sin la menor duda, y lo hubiesen aireado a los cuatro vientos. No solo no lo hicieron, sino que a pesar

del antagonismo existente entonces, pidieron humildemente al Gobierno de los EE.UU. que Armstrong y Aldrin llevaran hasta nuestro satélite aquel recuerdo entrañable de dos de sus cosmonautas fallecidos fatídicamente; 2) ¿y de dónde venían los datos que estuvimos recibiendo de la Luna en Fresnedillas y otras estaciones de la red, durante los diez años siguientes al *Apollo XI*, bajo la denominación de ALSEP? (*Apollo Lunar Surface Experiment Package*—Paquete de Experimentos de la Superficie Lunar el Apollo).

♫ ¡Y volver, volver, voolveeerrr...! ♫

El asesinado presidente Kennedy había prometido poner un hombre en la Luna, <<... y devolverlo sano y salvo a la Tierra...>>, y su antaño oponente político, y en ese momento vigente presidente Nixon, manifestó la misma esperanza en su mensaje telefónico a Armstrong y Aldrin:

<<...For one priceless moment, in the whole history of man, all the people on this earth are truly one. One in their pride in what you have done, and one in our prayers that you will return safely to Earth.>>

(... Por un momento inapreciable, en toda la Historia del Hombre, todos los pueblos de este planeta son realmente uno. Uno en su orgullo por lo que habéis hecho, y uno en sus plegarias para que **regreséis felizmente a la Tierra.**)



Fig.28. Bosquejo a carboncillo de Paul Calle de la inédita llamada telefónica del Presidente de los EE.UU. Richard Nixon, a los dos astronautas que se hallaban hollando nuestro satélite. (Julio 1969)

Quienes estuvimos metidos de lleno en el vuelo del *Apollo XI* en Fresnedillas (*Madrid Apollo Prime*) y en Robledo (*Madrid Apollo Wing*), sabíamos que tan arriesgado había sido llevar a su tripulación a la Luna, como conseguir después dejarlos aceptablemente intactos en la cubierta del portaviones *USS Hornet*, que surcaba el Pacífico.

Los cálculos de navegación para el retorno abocaron a que fuera de nuevo la estación de Fresnedillas la responsable (*Prime station*) del control total desde la Tierra de la secuencias definitiva para el éxito final de la misión *Apollo XI*: el despegue desde la superficie lunar de la parte superior del *Eagle*, que tenía que llevar a Armstrong y a Aldrin al encuentro con su colega Collins, tras su 26ª órbita alrededor de la Luna.

Los especialistas de Houston ya habían descubierto que la *Eagle* estaba a unos seis kilómetros del lugar previsto originalmente para el aterrizaje, y por lo tanto tuvieron que hacer correcciones en los datos de navegación almacenados en los ordenadores de las dos naves. Fresnedillas (*Apollo Prime*) transmitió los nuevos datos al *Eagle*, y Robledo (*Apollo Wing*) hizo otro tanto con el *Columbia*.

Pero la mermada reserva de oxígeno para Armstrong y Aldrin (sólo para un día más), decidió a los navegantes no demorar más la salida al encuentro de Collins. Dado el meridiano de Madrid, la NASA recabó la colaboración de las estaciones de Canarias, y las lejanas Ascensión y Tananarive (Madagascar), aunque sus reducidas antenas de 9 metros de diámetro limitaban seriamente su capacidad de transmisión y recepción para la enorme distancia Tierra – Luna de unos 380.000 km.

Sabíamos que el motor del cohete que les debería elevar desde la parte inferior del módulo lunar, no había sido probado nunca antes en la Luna, y las pruebas que se hicieron en la Tierra, fundieron las toberas del cohete. Pero uno de los ingenieros de la Grumman aseguró, llenando la pizarra de fórmulas, que teóricamente ese cohete daría el do de pecho en la Luna.

Aquella iba a ser una primicia con todas sus consecuencias, y si al pulsar el botón de encendido no había respuesta, la tragedia estaba servida. Sabemos que en el control de Houston no se oía una mosca en los minutos previos al intento de despegue, pero es que en Fresnedillas el silencio se masticaba. Y el cohete funcionó. Aunque la angustia previa que pasó la tripulación del *Eagle*, y quienes sabíamos en tierra de la dudosa eficacia del cohete, no se puede medir ni pesar.

Arriba, Collins se había confeccionado una chuleta con los 18 pasos más importantes que debía seguir al encontrarse con sus colegas de la *Eagle*, y para no extraviarla se la había colgado al cuello.



Fig.29. Traslado desde el portaviones USS *Hornet* a Houston del preciadísimo material traído de la Luna por la tripulación del *Apollo XI*, en su embalaje hermético (Mobil Quarentine Facility – MQF-).

El reencuentro de los tres camaradas fue suave, y tras los emocionados mini abrazos (no había sitio para grandes aspavientos), trasladaron con toda premura el material obtenido en la superficie lunar: 21,55 kg de rocas, fotos, películas, y el detector de viento solar. Se separaron del *Eagle* que tan bien había cumplido, y le enviaron a estrellarse contra la Luna, hecho que se pudo comprobar al recibir en tierra las bruscas señales del sismómetro instalado por Armstrong en el paquete EASEP.

A partir de ese momento se inició el ansiado retorno a casa acelerando desde 6.000 kilómetros por hora hasta

alcanzar los 40.000 km/h al horadar las capas altas de la atmósfera. Michael Collins aprovechó para tener un bonito detalle con quienes habíamos intervenido (y seguíamos haciéndolo) en el conjunto de la misión *Apollo XI*. Aprovechó una retransmisión televisada para la Tierra, y nos regaló los oídos con este mensaje televisado que todos pudimos ver y oír en Fresnedillas, Robledo y Maspalomas:

<<*This operation [Apollo XI] is somewhat like the periscope of a submarine. All you see is the three of us, but beneath the surface are thousands and thousands of others, and to all of those, I would like to say, **Thank you very much!***>>

(Esta operación [*Apollo XI*] es de alguna forma como el periscopio de un submarino. Todo lo que se ve es a nosotros tres, pero debajo de la superficie hay miles y miles de otros, y a todos esos otros, me gustaría decir: **¡Muchas gracias!**)

Tras aquella aleluya a nuestra labor colectiva internacional, el resto del vuelo continuó sin ninguna incidencia digna de mención, esperando nosotros aquí abajo que aquellos osados viajeros culminaran con bien su hazaña. Para ello, la tripulación del módulo de mando *Columbia* tuvo que separarse del fiel compañero llamado “módulo de servicio” (*Service Module*), antes de zambullirse en la peligrosa atmósfera.

El ángulo de reentrada había sido prefijado para caer sobre el Océano Pacífico, por lo que Fresnedillas se despidió de su participación activa por falta de visibilidad. No obstante, nadie se movió de la Estación hasta que pudimos ver en los monitores de televisión el chapoteo del reducido módulo de mando (*Command Module – Columbia*), aún sujeto a los paracaídas. En ese momento los relojes GET (*Ground Elapsed Time*) de nuestras salas de Operaciones marcaban 195 horas 18 minutos y 35 segundos desde que despegaron de Cabo Cañaveral el pasado 16 de julio. Y esa fue la hora en que se desbordó la emoción y el vocerío en todos y cada uno de los centros de la NASA que habíamos vivido, sufrido y saboreado aquella gran hazaña.



Fig.30. Los héroes del **Apollo XI** visitan la capital de España y son aclamados por el público que se aglomera en la Gran Vía madrileña. (7/10/1969)

Tres meses después, en octubre de 1969, la tripulación del *Apollo XI* vino a Madrid en visita oficial. No me privo de reproducir aquí las palabras que Neil Armstrong en nombre propio y de sus compañeros, dijo en la recepción de la Embajada norteamericana, ante una apelotonada masa de periodistas e invitados:

<<Sin las vitales comunicaciones mantenidas entre el Apolo XI y la estación Apolo [Fresnedillas], en Madrid, podemos afirmar que nuestro aterrizaje en la Luna no hubiera sido posible. >>

¡Bien dicho!

Splashdown party.

La prueba de fuego de la Estación de Seguimiento Espacial de Fresnedillas se la debemos al renombrado *Apollo VIII*, que llevó a tres hombres a circunvalar la Luna por primera vez en la Historia. Aquel hito ocurrió en las Navidades de 1968, y estableció una costumbre en la Estación que llegó a hacer tradición, la de celebrar el feliz retorno de los astronautas, cada vez que éstos eran recuperados en el Océano Pacífico procedentes de la Luna.

Cuando el *Mission Space Center* (Centro Espacial de Houston) llamaba a *Madrid Apollo* (Fresnedillas) para felicitarnos -y felicitarse-, por el éxito de la misión, nos despedía con un: *<<Madrid, thank you for your excellent support! You are released! >>*, que venía a ser el *¡Rompan filas!* de la fenecida Mili. La traducción en definitiva era: *¡Gracias por vuestro excelente apoyo! ¡Os podéis ir!*

Se arriaba la bandera del Programa *Apollo*, que había sido izada a los pocos minutos del despegue del enorme cohete *Saturno V*, en el mástil delante del edificio de Entrenamiento, procedíamos a recoger y retirar los equipos auxiliares de medida, y a archivar la montonera de libros, manuales y *twixes* (mensajes por teletipo), que cubrían las mesas y encimeras.

Cuando las salas de Operaciones (*USB, Telemetry, Computers y Communications*) quedaban en perfecto orden de revista, asaltábamos los coches (entonces Seat 124), y nos dirigíamos formando un convoy de una treintena de vehículos, a disfrutar el merecido *Splashdown Party* (Fiesta del Amaraje), en un conocido restaurante de Valdemorillo.

Como el acontecimiento a celebrar era muy especial, casi todos nos poníamos ese día corbata, para añadir solemnidad al acto. A uno de los jefes norteamericanos se le ocurrió ponerse a la puerta del recinto, con unas enormes tijeras de sastre, y rebanar las corbatas una por una, según íbamos accediendo al local. Los muñones de corbatas, de amplia gama de diseños y colores, eran claveteados después en una galería de madera que se izaba

orgullosamente sobre la ventana que daba a la Plaza Mayor, desde donde la podían ver todos los viandantes. Aquel atípico cortinaje de corbatas permanecía intacto hasta el siguiente *splashdown party*, -normalmente seis meses después-, en que una nueva remesa se claveteaba sobre la anterior, y así año tras año hasta diciembre de 1972, en que el *Apollo XVII* cerró el ciclo de viajes a nuestro satélite natural.

Es una pena que no se haya conservado aquel atípico souvenir, que ahora podría albergarse con todo honor en el Museo Lunar de Fresnedillas de la Oliva, junto a tantas otras reliquias de aquellos tiempos irrepetibles.



Fig.31. Amena entrevista del autor al astronauta Ronald Evans, piloto del módulo de mando del **Apollo XVII**. (1976)

La Estación de Seguimiento de Vuelos Espaciales Tripulados de Fresnedillas, ya totalmente en manos de profesionales españoles, volvió a protagonizar en primera persona el último lanzamiento lunar tripulado, el *Apollo XVII*, (7/dic/1972), y la última despedida del hombre de la vieja Selene (14/dic/1972), por muchos, muchos años.

En 1976 tuve el honor y el placer de trabajar en Fresnedillas con el astronauta Ronald E. Evans, piloto del módulo de mando (CSM- *America*), quien me regaló una copia de la placa metálica que dejó firmada la tripulación (Eugene Cernan, Ronald Evans y Harrison Schmitt), sujeta a una de las patas de su módulo lunar (LM - *Challenger*). Dice así:

<<Here man completed his first explorations on the Moon. December 1972. A.D. May the spirit of peace in which we came be reflected in the lives of all mankind.>>

(Aquí el hombre ha completado sus primeras exploraciones lunares. Diciembre 1972. Era Cristiana. Que el espíritu de paz con el que vinimos se vea reflejado en las vidas de toda la Humanidad.)



Fig.32. Copia de la placa del **Apollo XVII**, que regaló al autor el piloto del módulo de mando *América*, Ron Evans, cuyo original permanece fijado a una de las patas del módulo lunar **Challenger**. (1976).

Al cerrar estas líneas con un reducido salpicón de recuerdos personales, limitado por el espacio disponible, estamos a las puertas de que se cumpla el medio siglo de la mayor hazaña colectiva jamás conseguida por el ser humano. Si de alguna forma he abierto los ojos a mis compatriotas sobre la decisiva participación que allá entre 1968 y 1972 tuvieron un equipo de españoles, que a requerimiento de la NASA se hicieron acreedores de la indescriptible responsabilidad de poner sobre la superficie lunar a doce hombres, y traerlos de vuelta a su planeta hogar sanos y salvos, entonces me daré por satisfecho. Envío un abrazo a mis colegas de Fresnedillas, Robledo y Maspalomas, fallecidos o vivos, recordándoles que todos juntos hicimos Historia, y que por lo tanto es nuestra Historia.



*Fig.33. Histórica foto de toda la plantilla del personal –ya íntegramente español–de la Estación de Seguimiento Espacial de Fresnedillas, al concluir el legendario Programa **Apollo**. (1972)*

José Manuel Grandela Durán
Ex Ingeniero Controlador de Naves Espaciales (INTA-NASA).

Capítulo 2

La tabla periódica: algo más que un icono de la ciencia.

La tabla periódica (TP) es uno de los iconos más relevantes y universales de la ciencia moderna. Probablemente cualquier persona que haya cursado educación secundaria es capaz de recordar que constituye una clasificación de los elementos químicos. A veces, se rememora con cierto desagrado, como un ejercicio de memorización estéril acompañado de unos “misteriosos” números denominados “valencias”. Para quien ha profundizado algo más, en bachillerato, en estudios universitarios, o con lecturas sobre temas científicos, es fácil que la identifique como una herramienta de gran utilidad y con una historia curiosa para su construcción. En este capítulo se muestra brevemente qué significa, por qué es importante y cómo sirve para ordenar los elementos químicos de los que está hecho todo: el aire, nuestro cuerpo, las flores... Una idea que se intenta transmitir es que es mucho más que un icono; es un ejemplo de cómo se desarrolla la ciencia, fruto de una labor colectiva de personas de muy diversos países y generaciones. Además, se explicará que la construcción de la TP y sus aplicaciones no incumben únicamente a la química, sino también a otras ciencias. Por otra parte, es un logro científico donde la labor del profesorado de las distintas etapas educativas ha sido primordial.

Celebración en 2019 del Año Internacional de la Tabla Periódica de los Elementos Químicos.

En diciembre de 2017 la *Asamblea General de Naciones Unidas* proclamó 2019 como *Año Internacional de la Tabla Periódica de los Elementos Químicos*, culminando así una iniciativa promovida a través de la UNESCO y de Rusia. La idea surgió del investigador, profesor y divulgador inglés Martyn Polyakoff (de madre inglesa y padre ruso), que se lo sugirió a la profesora rusa Natalia Tarasova, presidenta de la *Unión Internacional de Química Pura y Aplicada* (IUPAC). El objetivo es conmemorar el 150º aniversario de la propuesta de ordenación de los elementos químicos conocidos entonces, en forma de tabla, por el ruso Dimitri Ivánovich Mendeléiev, así como el 100º aniversario de la creación de la propia IUPAC. Con la efeméride, se reconoce la importancia de las ciencias básicas en el desarrollo sostenible y en la solución de problemas globales (energía, educación, agricultura, salud...).

Durante todo el año 2019 se están llevando a cabo, a nivel mundial, multitud de acciones y actividades en torno a este tema [1]. Por ejemplo, en España, se ha emitido un sello que complementa al diseñado en 2007 por el

centenario de la muerte de Mendeléiev [2]. Ambos sellos se recogen en la Fig. 1.

Los elementos químicos.

La noción de elemento químico es un aspecto central de la química y para entender la utilidad de la TP. Para algunos filósofos presocráticos todos los cuerpos estaban compuestos por cuatro constituyentes fundamentales (tierra, aire, fuego y agua), en distinta proporción, a los que Aristóteles añadiría un “quinto elemento” muy sutil, el éter, formador de la esfera celeste, que aún se rememora cuando se habla de la “quintaesencia” (lo más puro, fino y acendrado de una cosa, según la RAE). Clasificaciones análogas forman parte de la tradición de las grandes culturas asiáticas. En

el siglo XVI, el alquimista y médico suizo Paracelso propuso que había tres “principios fundamentales” (sal, mercurio y azufre) que hoy en día se pueden entender como propiedades (solidez, fluidez y combustibilidad), más que como ingredientes fundamentales de la materia². En todo caso, es en el siglo XVII cuando el anglo-irlandés Robert Boyle establece que la materia está formada por “ciertos cuerpos primitivos y simples” que componen los “cuerpos mixtos”, en su obra *The Sceptical Chymist* (1661), y propone que un elemento (“cuerpo simple”) es una sustancia que no puede reducirse (“ser analizada”) a otra más sencilla. El francés Antoine-Laurent de Lavoisier, en su *Traité Élémentaire de Chimie* (1789), reforzó este concepto, postulando que “sustancia simple” es la que no se puede descomponer en otras por procedimientos químicos. Algunas de sus sustancias (en total 33), como la luz y el calórico hoy no se consideran elementos y otras, como la cal, son compuestos.

El inglés John Dalton, que con el tiempo serviría para denominar a la alteración en la capacidad de distinguir los colores (daltonismo) que estudió en él mismo, recuperó la teoría atomista de los griegos para explicar la composición de la materia. En 1808 explicó en *A New System of Chemical Philosophy* que las propiedades específicas de los elementos derivan de sus átomos constituyentes, que se podían visualizar como pequeñas esferas. Su teoría sugería que los átomos de distintos elementos poseen diferente masa y que se combinan para formar compuestos en proporciones de números enteros sencillos (1:1, 1:2, 2:3...). Fue de los primeros en determinar los pesos atómicos, labor que perfeccionó el sueco Jöns Jacob Berzelius y, de forma más concluyente, el italiano Stanislao Cannizzaro, quien distinguió entre pesos



Fig.1. Sellos sobre la tabla periódica (imágenes cedidas por la S.E. Correos y Telégrafos S.A.).

²Paracelso es el pseudónimo que eligió para sí mismo Theophrastus Bombast von Hohenheim, por considerarse “semejante a Celso” (autor romano de libros de medicina del siglo I). Tuvo mucha fama porque se pensaba que consiguió la “transmutación” del plomo en oro. Destacó en toxicología, donde es conocida su sentencia “*dosis sola facit venenum*” (la dosis hace al veneno).

moleculares y atómicos³ y que protagonizaría el Congreso de Karlsruhe (Alemania), reunión internacional de química, celebrada en 1860, donde se puso de manifiesto que existían ciertas regularidades en las propiedades de los elementos químicos. Desde entonces, hasta el primer tercio del siglo XX, se fue descubriendo cada vez más sobre la constitución de los átomos y se entendió que consisten en pequeños núcleos formados por protones (carga positiva) y neutrones (carga neutra) rodeados de electrones (carga negativa y de mucha menor masa) [3]. La idea actual de elemento químico es que se trata de la sustancia que no se puede descomponer en otras más simples por métodos químicos y que está formada por átomos del mismo tipo que, al combinarse entre sí, forman compuestos [4].

La construcción de la tabla periódica: introducción histórica.

Para simplificar y racionalizar el estudio de la química de los elementos químicos, hubo múltiples intentos de clasificarlos en función de sus propiedades y pesos atómicos a lo largo del siglo XIX [5-8]. Son emblemáticas las “triadas” del alemán Johann Wolfgang Döbereiner, que clasificó en 1829 los elementos, por propiedades químicas similares, en grupos de tres (como calcio, estroncio y bario) y advirtió que, en cada *triada*, el peso atómico de uno era la media aritmética de los otros dos aproximadamente. El geólogo francés Alexandre Émile Béguyer de Chancourtois, en 1862, colocó los elementos conocidos (y algunos compuestos) por orden creciente de peso atómico, sobre una curva helicoidal inscrita alrededor de un cilindro vertical con una circunferencia de 16 unidades (peso atómico del oxígeno), en la que los elementos que caían en la misma vertical tenían propiedades semejantes. Sugirió una ordenación que denominó *vis tellurique* (tornillo telúrico), por estar el telurio en el centro de la gráfica. Para Scerri, aunque Chancourtois recibió poca atención en su tiempo, fue quien realmente descubrió la periodicidad química [6].

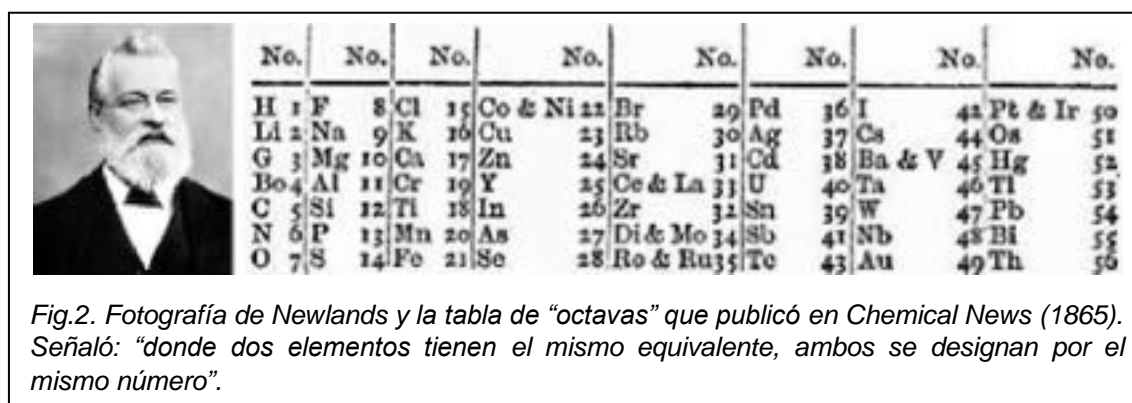


Fig.2. Fotografía de Newlands y la tabla de “octavas” que publicó en *Chemical News* (1865). Señaló: “donde dos elementos tienen el mismo equivalente, ambos se designan por el mismo número”.

³El peso atómico de la época (inicialmente referido por el concepto químico, hoy en desuso, de “equivalente”) se conoce actualmente como *masa atómica relativa*. Las diferencias sobre el significado de este tipo de conceptos conllevan cierta controversia. Otro ejemplo del cuidado que hay que prestar al significado y evolución histórica de conceptos es la diferenciación entre valencia (término introducido a finales del siglo XIX para designar el “poder de combinación” de un elemento), carga formal y número de oxidación.

En 1865, el químico inglés John Alexander Reina Newlands⁴ apreció que cuando los elementos se ordenaban por pesos atómicos crecientes, cada ocho elementos existían semejanzas en las propiedades físicas y químicas (Fig. 2). Como señaló en ello cierto parecido con la escala musical, propuso que se podría denominar, “de forma provisional”, como “Ley de las octavas” que fue ridiculizada pero, años después, fue reconocida.

De forma independiente, dos asistentes al Congreso de Karlsruhe, Mendeléiev y el alemán Julius Lothar Meyer (menos conocido en su tiempo), establecieron que existe una ley periódica, por la que las propiedades de los elementos (unos 60 entonces) se repiten con cierta periodicidad al aumentar su peso atómico. La tabla de Meyer está fechada en diciembre de 1868, pero no se publicó hasta 1870, mientras que la de Mendeléiev fue comunicada por Menshutkin (por enfermedad del autor), en octubre de 1868, a las Sociedad Rusa de Química y se publicó en 1869 (Fig. 3). Mendeléiev fue más consciente de la importancia de su clasificación, que fue modificando durante las cuatro décadas siguientes.

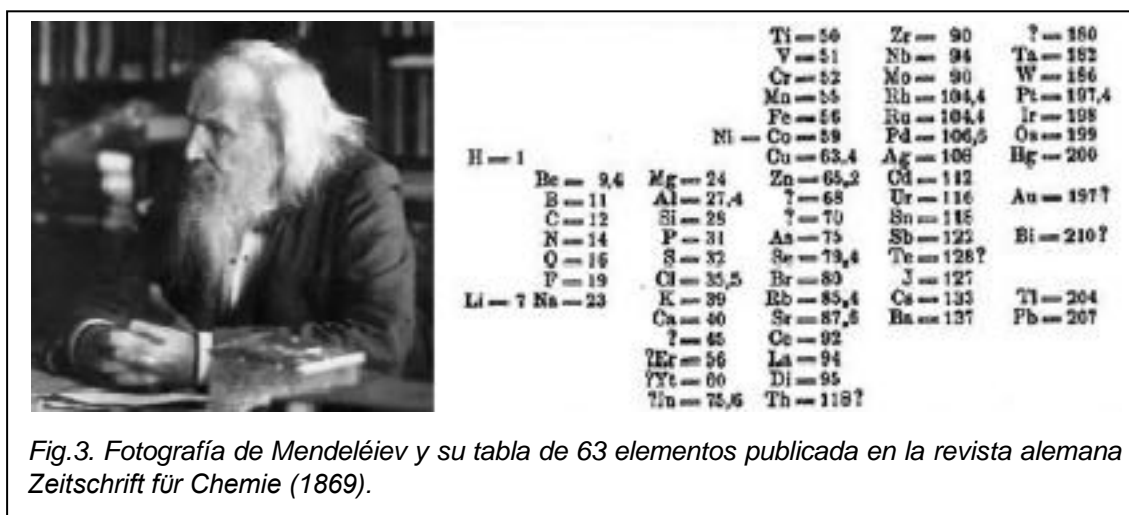


Fig.3. Fotografía de Mendeléiev y su tabla de 63 elementos publicada en la revista alemana Zeitschrift für Chemie (1869).

Visto siglo y medio después resulta admirable la intuición de Mendeléiev, considerando la inexactitud de muchos de los datos de que disponía. De su tabla se deducía una periodicidad de las propiedades, referidas esencialmente a las proporciones con las que se combinaban los elementos, y que dependían, a su juicio, de los pesos atómicos. Lo más relevante es que predijo la existencia de elementos que se descubrirían después, como el galio (*ekaaluminio*) y el germanio (*ekasilicio*), y que cambiara el orden en algún par de elementos [9].

Gracias a la TP de Mendeléiev y otras como las comentadas, se pudieron relacionar las propiedades de los elementos y sus compuestos con otros de comportamiento similar, lo que sirvió de impulso para la racionalización de la química. Su tabla supuso un verdadero renacimiento de esta ciencia y se considera que varios descubrimientos posteriores tienen su base en la búsqueda de las causas de las irregularidades que señaló. Entre los

⁴El apellido Reina se debe al origen italiano de su madre. Newlands participó como voluntario con Giuseppe Garibaldi en la conquista del reino de Nápoles (1860) para la reunificación italiana.

problemas estaban, como se ha comentado, los “huecos” de elementos y algunas inversiones de orden, como telurio-yodo, que en un principio justificó por errores en el cálculo de pesos atómicos. Estas inversiones se explicarían en 1913, como otros aspectos esenciales de la TP, por el físico inglés Henry Gwyn Jeffreys Moseley⁵, al justificar el concepto de número atómico, mediante



Fig.4. Moseley en el laboratorio Balliol-Trinity de la Universidad de Oxford (1910) y la representación de la raíz cuadrada de las frecuencias de emisión de rayos X de los elementos en relación a su número atómico, publicada en *Philosophical Magazine* (1914).

la ley que lleva su nombre, que establece la relación empírica entre la longitud de onda de rayos X emitidos por los átomos, con su número atómico (Fig. 4). Fue la base para establecer la “ley periódica” moderna, según la cual, las propiedades físicas y químicas de los elementos muestran tendencias periódicas al ponerlos en orden por número atómico (y no del peso atómico, como la primitiva ley periódica de Mendeléiev).

Las aportaciones de Mendeléiev, Meyer y otros científicos del siglo XIX, en la construcción de la TP, se acrecientan si se considera que, en su época, no se conocía la estructura de la materia a nivel atómico, y que la separación y la purificación de los elementos eran rudimentarias. Estos autores se solían referir a un “sistema periódico”, término que permanece aún en idiomas como el alemán (*periodensystem*) si bien, con el tiempo, se fue denominando “tabla periódica” en otros como inglés (*periodic table*), francés (*tableau o table périodique*) y español (donde se emplean ambas denominaciones), al tratarse de una ordenación de los elementos en filas y columnas.

Entre otros ejemplos de investigadores que han destacado en la historia de la construcción de la TP, se citan aquí algunos más, por su especial importancia. En 1868 se

descubrió el primer gas noble, helio, por las líneas espectrales de la cromosfera del Sol. Supuso una gran contrariedad para Mendeléiev, pues no había previsto

⁵Moseley falleció con solo 27 años en la batalla de Galípoli (Turquía, 1915) en la Primera Guerra Mundial. Su aportación fue fundamental para la comprensión de la estructura atómica y de la TP.

un lugar adecuado en su TP. En 1895 el físico inglés Lord Rayleigh descubrió el argón, en colaboración con el químico escocés William Ramsay (Fig. 5). Este último descubrió tres años después otros tres gases nobles (Kr, Ne y Xe). Finalmente, la existencia de los gases nobles (también denominados inertes) fue una brillante confirmación de la ley periódica y Mendeléiev los incluiría como grupo 0 en su TP. El químico y físico inglés Frederick Soddy explicó la importancia del peso atómico ponderado (teniendo en cuenta la proporción de cada isótopo) y fue quien acuñó el término “isótopo”⁶ [10]. El físico danés Niels Henrik David Bohr realizó contribuciones fundamentales para la comprensión de la estructura del átomo y de la mecánica cuántica, como el físico alemán Erwin Schrödinger. Más recientemente, a partir de los años cuarenta del pasado siglo, el químico nuclear norteamericano Glenn T. Seaborg colaboró activamente en el descubrimiento y aislamiento de diez elementos transuránicos (de mayor número atómico que el uranio) que se añadieron a la TP y son: Pu, Am, Cm, Bk, Cf, Es, Fm, Md, No y Sg (este se denominó seaborgio en su honor). Además, sugirió que tanto los lantánidos como los actínidos iban completando los subniveles electrónicos “f” y no los “d” como se pensaba.

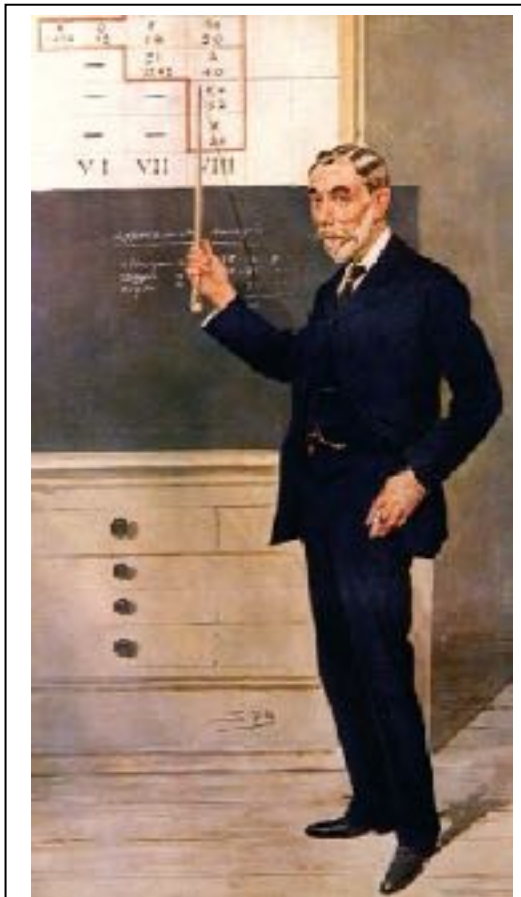


Fig.5. Litografía de Ramsay realizada por el conocido caricaturista Leslie Ward, titulada “Chemistry” (Vanity Fair, 1908).

En la Fig. 6 se recoge un resumen de las teorías atómicas más relevantes, realizado por el profesor Andy Brunning, muy conocido en los últimos años por sus infografías sobre conceptos relacionados con la química [11]. La mecánica cuántica, consolidada a finales de la década de los 20, aportaría la explicación actual de la TP. Así, las propiedades de los elementos varían periódicamente con el número atómico (número de protones del núcleo del átomo, coincidente con su número de electrones) que, como se ha indicado, no es proporcional al peso atómico (dependiente del número de protones pero también del número de neutrones presentes en el núcleo atómico) en todos los casos. Los electrones se distribuyen en orbitales atómicos con energía “cuantizada”, no continua. La periodicidad no se produce para los mismos intervalos, sino sucesivamente, cada 2, 8, 8, 18, 18, 32 y 32 unidades de número atómico. Las propiedades se repiten aproximadamente al repetirse la disposición de los electrones de la capa electrónica más externa (de valencia). Así, los alcalinos, como el sodio y el potasio, poseen propiedades

⁶Creó un neologismo basado en las palabras griegas *isos* (igual) y *topos* (lugar), refiriéndose a átomos con igual número de protones (número atómico) y distinta masa atómica (por poseer diferente número de neutrones). Explicó que los denominó isótopos o elementos isotópicos, por ocupar el mismo lugar en la TP y ser químicamente idénticos.

análogas, como la formación de cationes M^+ , porque la configuración de su capa de valencia es del tipo s^1 , mientras que los halógenos, como el flúor y el cloro, tienen propiedades similares, como la formación de aniones X^- , porque su configuración es del tipo s^2p^5 , o todos los gases nobles son muy poco reactivos porque completan su capa de valencia (s^2p^6).

Esto no es sencillo, pero su potencialidad didáctica es tan relevante que en muchos países, como España, se aborda (desde hace décadas) en los currículos oficiales para alumnos de unos 13 años de edad; en otros, se opta por enseñar una química más macroscópica en la etapa introductoria y estos aspectos se tratan en cursos superiores.

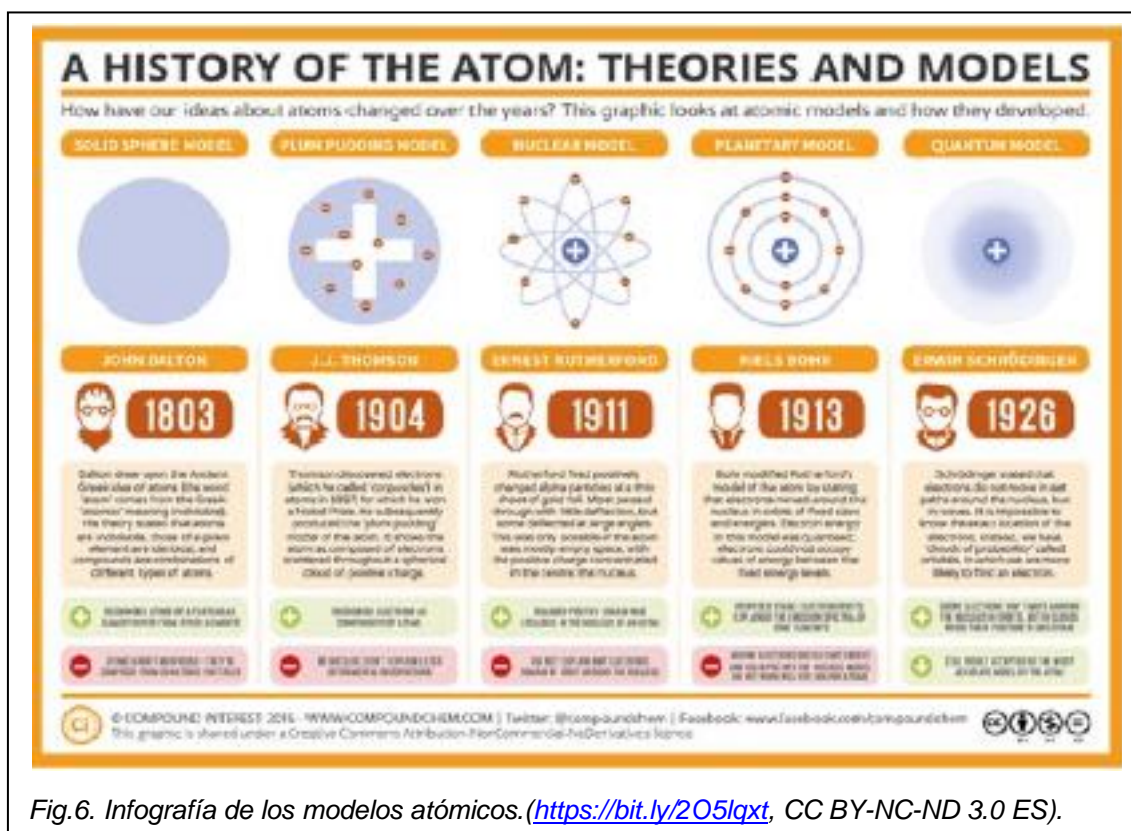


Fig.6. Infografía de los modelos atómicos. (<https://bit.ly/2O5lqxt>, CC BY-NC-ND 3.0 ES).

Formas de representar la tabla periódica.

La labor del profesorado y de las editoriales de libros de texto en la evolución de la TP ha sido crucial. El propio Mendeléiev se esforzó en clasificar los elementos químicos con objeto de racionalizar el estudio de la química inorgánica en un libro de texto. Parece ser que era muy aficionado a los juegos de cartas de solitarios, con los que se entretenía en sus frecuentes viajes en tren. Durante uno de estos, para visitar varias fábricas, preparó una carta para cada elemento, donde anotó en cada una el peso atómico y propiedades típicas; al ordenarlas por peso atómico, se convenció de que las propiedades dependían periódicamente de los pesos atómicos.

En la Fig. 7 se muestra la considerada como la TP educativa más antigua conservada; se encontró de casualidad hace poco, en la Universidad de Saint Andrews (Escocia) y se supone que data de alrededor de 1885. En la Fig. 8 se muestra otro ejemplo de TP académica muy común en los años 30 en España, donde se seguía el modelo propuesto en 1926 por el químico estonio Andreas von Antropoff⁷. Aparte de la recogida en esta figura, conservada en el I.E.S. San Isidro de Madrid, hay otra del tipo von Antropoff en formato mural en la Universidad de Barcelona. Cada grupo se identifica con diferentes colores (elegidos por alguna propiedad de los elementos correspondientes), mediante barras de unión entrelazadas [12].



Fig. 7. Tabla periódica de alrededor de 1885.



Fig. 8. Tabla periódica según el modelo de von Antropoff conservada en el I.E.S. San Isidro de Madrid. (Fotografía realizada y remitida por las profesoras M^a. José Gómez y Rosario Ramos, respectivamente).

Se han ideado diferentes estructuras a lo largo de la historia a partir de la TP inicial de Mendeléiev, colocando los elementos de formas variadas. Las familias de elementos con propiedades similares, que en un principio se representaban en filas, ahora (como ya hizo Mendeléiev en sucesivas actualizaciones) se recogen en columnas. La forma más conocida y usada, que aparece en casi todos los libros de texto y publicaciones actuales, es la alargada que se muestra en la Fig. 9. En este modelo hay 7 periodos. A veces se incluye un octavo con solo dos elementos, Uue (119) y Ubn (120), aún no descubiertos, y 18 columnas numeradas de la 1 a la 18. Hace unos años, las columnas (familias de elementos) se designaban por grupos IA a VIII A (elementos representativos) y de IB a VIIB, más un VIII (que incluye los grupos del Ni, Fe y Co), para agrupar los metales de transición. Algunos grupos tienen denominación especial, como gases nobles, alcalinos, alcalinotérreos y halógenos.

⁷Doctorado en Heidelberg (Alemania), se formó con Ramsay en Londres y fue profesor en San Petersburgo. Tras ser arrestado en 1918 por el Soviet de Petrogrado, pasó a trabajar en la Universidad de Bonn, donde llegó a decano y destacó por ser un nacionalsocialista convencido.

Existen aún ciertas discrepancias sobre algunos aspectos. Así, en algún caso se coloca el He encima del Be, por ser la columna de las configuraciones electrónicas s^2 , pero se suele ubicar encima del Ne, como típico gas noble que es. Con el H ocurre algo similar; aunque se suele colocar encima del Li, porque ambos son de configuración s^1 , a veces se dispone encima del F por presentar propiedades más similares a las de los halógenos que a las de los metales alcalinos. Existen muchas variedades de colores, que suelen representar tres grandes familias de elementos: metales, no metales y semimetales (también denominados metaloides). Pero también se usan colores, a veces, para distinguir los elementos que se presentan como sólido, líquido o gas, en condiciones especificadas de temperatura y presión, así como si se trata de elementos sintéticos.

IUPAC Periodic Table of the Elements

| | | | | | | | | | | | | | | | | | |
|------------------------------------|----------------------------------|---------------------------------------|-----------------------------------|------------------------------------|------------------------------------|------------------------------------|----------------------------------|-----------------------------------|--------------------------------------|-------------------------------------|----------------------------------|--------------------------------------|------------------------------------|-------------------------------------|---------------------------------|----------------------------------|--------------------------------|
| 1 H Hydrogen 1.00794 | | | | | | | | | | | | | | | | | 18 He Helium 4.002602 |
| 3 Li Lithium 6.941 | 4 Be Beryllium 9.012182 | | | | | | | | | | | 5 B Boron 10.811 | 6 C Carbon 12.011 | 7 N Nitrogen 14.006434 | 8 O Oxygen 15.999 | 9 F Fluorine 18.9984032 | 10 Ne Neon 20.1797 |
| 11 Na Sodium 22.98976928 | 12 Mg Magnesium 24.304 | | | | | | | | | | | 13 Al Aluminum 26.9815386 | 14 Si Silicon 28.0855 | 15 P Phosphorus 30.973762 | 16 S Sulfur 32.06 | 17 Cl Chlorine 35.45 | 18 Ar Argon 39.948 |
| 19 K Potassium 39.0983 | 20 Ca Calcium 40.078 | 21 Sc Scandium 44.955912 | 22 Ti Titanium 47.88 | 23 V Vanadium 50.9415 | 24 Cr Chromium 51.9961 | 25 Mn Manganese 54.938045 | 26 Fe Iron 55.845 | 27 Co Cobalt 58.933195 | 28 Ni Nickel 58.6934 | 29 Cu Copper 63.546 | 30 Zn Zinc 65.38 | 31 Ga Gallium 69.723 | 32 Ge Germanium 72.630 | 33 As Arsenic 74.9216 | 34 Se Selenium 78.96 | 35 Br Bromine 79.904 | 36 Kr Krypton 83.80 |
| 37 Rb Rubidium 85.4678 | 38 Sr Strontium 87.62 | 39 Y Yttrium 88.90584 | 40 Zr Zirconium 91.224 | 41 Nb Niobium 92.90638 | 42 Mo Molybdenum 95.94 | 43 Tc Technetium 98.9062 | 44 Ru Ruthenium 101.07 | 45 Rh Rhodium 102.9055 | 46 Pd Palladium 106.36 | 47 Ag Silver 107.8682 | 48 Cd Cadmium 112.411 | 49 In Indium 114.818 | 50 Sn Tin 118.710 | 51 Sb Antimony 121.757 | 52 Te Tellurium 127.6 | 53 I Iodine 126.905 | 54 Xe Xenon 131.29 |
| 55 Cs Cesium 132.90545196 | 56 Ba Barium 137.327 | 57-71 Lanthanides | 72 Hf Hafnium 178.49 | 73 Ta Tantalum 180.94788 | 74 W Tungsten 183.84 | 75 Re Rhenium 186.207 | 76 Os Osmium 190.23 | 77 Ir Iridium 192.222 | 78 Pt Platinum 195.084 | 79 Au Gold 196.966569 | 80 Hg Mercury 200.59 | 81 Tl Thallium 204.3833 | 82 Pb Lead 207.2 | 83 Bi Bismuth 208.9804 | 84 Po Polonium 209 | 85 At Astatine 210 | 86 Rn Radon 222 |
| 87 Fr Francium 223 | 88 Ra Radium 226 | 89-103 Actinides | 104 Rf Rutherfordium 261 | 105 Db Dubnium 262 | 106 Sg Seaborgium 263 | 107 Bh Bohrium 264 | 108 Hs Hassium 265 | 109 Mt Meitnerium 266 | 110 Ds Darmstadtium 267 | 111 Rg Roentgenium 268 | 112 Cn Copernicium 269 | 113 Nh Nihonium 270 | 114 Fl Flerovium 271 | 115 Mc Moscovium 272 | 116 Lv Livermorium 273 | 117 Ts Tennessine 274 | 118 Og Oganesson 276 |
| 57 La Lanthanum 138.90547 | 58 Ce Cerium 140.12 | 59 Pr Praseodymium 140.90766 | 60 Nd Neodymium 144.242 | 61 Pm Promethium 144.9126 | 62 Sm Samarium 150.36 | 63 Eu Europium 151.964 | 64 Gd Gadolinium 157.25 | 65 Tb Terbium 158.92532 | 66 Dy Dysprosium 162.50015 | 67 Ho Holmium 164.93032 | 68 Er Erbium 167.259 | 69 Tm Thulium 168.93032 | 70 Yb Ytterbium 173.05448 | 71 Lu Lutetium 174.967 | | | |
| 89 Ac Actinium 227 | 90 Th Thorium 232.0377 | 91 Pa Protactinium 231.03688 | 92 U Uranium 238.02891 | 93 Np Neptunium 237.04817 | 94 Pu Plutonium 244.06422 | 95 Am Americium 243.06136 | 96 Cm Curium 247.07647 | 97 Bk Berkelium 247.0703 | 98 Cf Californium 251.07958 | 99 Es Einsteinium 252.0833 | 100 Fm Fermium 257.1035 | 101 Md Mendelevium 258.1051 | 102 No Nobelium 259.1035 | 103 Lr Lawrencium 260.1035 | | | |

Fig.9. Tabla periódica de la IUPAC actualizada el 1 de diciembre de 2018. (<https://iupac.org/>).

El modelo de TP de 18 columnas se atribuye al químico norteamericano Groves Deming, que lo publicó en 1923. La empresa química y farmacéutica Merck la distribuyó ampliamente por centros de enseñanza media de Estados Unidos desde 1928 y, en la década siguiente, se difundió en manuales y enciclopedias de química por la empresa de material educativo Sargent-Welch.

Lo mínimo que suele figurar en la casilla de cada elemento en las TP actuales es su símbolo, nombre, número atómico y peso atómico, pero también se pueden incluir otras muchas propiedades (electronegatividad, puntos de fusión y ebullición, configuración electrónica, etc.).

Por no realizar figuras de TP excesivamente anchas, los elementos de transición interna (lantánidos y actínidos) se colocan generalmente en filas debajo de la TP general, con varias versiones. En la mostrada en la Fig. 9, se han colocado el La y el Ac (como primeros lantánidos y actínidos) debajo del Y.

En otras, se disponen en esas posiciones al Lu y al Lr (como últimos lantánidos y actínidos, respectivamente). También existen versiones de TP donde lantánidos y actínidos están insertados en la propia tabla principal, con lo que queda muy alargada. La interpretación de estos ordenamientos está de acuerdo a las configuraciones electrónicas de los elementos, de forma que, una vez se domina la cuestión, es muy fácil entender qué configuración electrónica posee cada elemento, solo observando dónde está ubicado en la TP.

En todo caso, ha habido, y sigue habiendo, multitud de propuestas menos convencionales, mostrando muchas veces importantes dotes de creatividad. Entre otras, se recogen dos ejemplos al respecto en la Fig. 10.

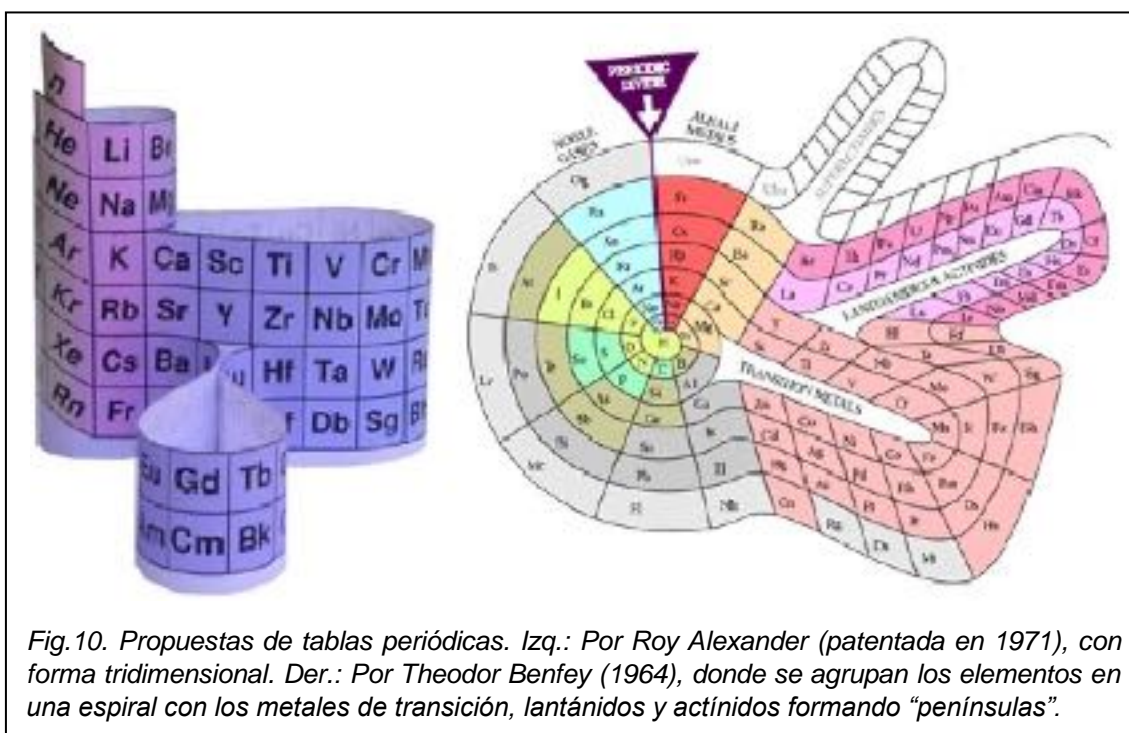


Fig.10. Propuestas de tablas periódicas. Izq.: Por Roy Alexander (patentada en 1971), con forma tridimensional. Der.: Por Theodor Benfey (1964), donde se agrupan los elementos en una espiral con los metales de transición, lantánidos y actínidos formando “penínsulas”.

Significado e interpretación de la tabla periódica moderna.

Como se ha indicado, podemos considerar la TP hoy en día como un reflejo de los resultados de la física cuántica en cuanto a la interpretación de las configuraciones electrónicas. Pero no conviene olvidar que también nos explica con bastante precisión por qué las propiedades atómicas de los elementos varían de forma periódica con el número atómico y, por tanto, con la posición que ocupan los elementos en la TP. Dicho de otro modo, gracias a los estudios de Mendeléiev y otros autores sobre cómo variaban las propiedades de los elementos, y gracias a los avances sobre la comprensión de la estructura atómica, se ha podido elaborar una TP de los elementos que nos permite razonar e interpretar cómo varían muchas propiedades. Dado que el estudio de esto en detalle excede los objetivos de este trabajo, se recomienda al lector interesado que visite la página Web de la *Royal Society of Chemistry* [13], que incluye también información para descargar como *app*. Se incluye allí una TP interactiva donde se puede elegir una presentación convencional de los

elementos (con sus símbolos) o con imágenes. En ella, se puede visualizar el estado físico de cada elemento (entre 0K y 6000K de temperatura), se incluyen una breve historia (incluso en formato *podcast* y en vídeo) e información sobre el origen etimológico de cada elemento y, quizá lo más interesante, se puede elegir una propiedad determinada (densidad, radio atómico, electronegatividad, punto de fusión, punto de ebullición y energía de la primera ionización) para apreciar de forma muy visual lo que significa la periodicidad.

Otra fuente de información muy relevante sobre la TP es la que se ofrece en la Web con el título de *Webelements* [14], que incluye incluso la posibilidad de comprar productos curiosos (modelos moleculares, camisetas, accesorios, pósteres, etc. sobre el tema.

Como ejercicio curioso, para ver si se conocen las posiciones de los elementos de la TP, se sugiere que el lector advierta si identifica qué representan los símbolos incluidos en la Fig. 11, donde se muestra una TP en chino. Otro ejercicio que se recomienda es buscar las biografías de los personajes (y otros relacionados con ellos) que aparecen en este capítulo e indagar sobre el origen de los nombres de los elementos, que remiten a localidades, países, científicos, propiedades, mitología, etc. [15].

Saber sobre la TP no es una mera curiosidad; conociéndola bien y entendiendo acerca de las variaciones periódicas que refleja, se puede justificar perfectamente cómo se forma un semiconductor tipo “n” o “p”, por “dopado” de un elemento como el Si o el Ge, se

pueden valorar alternativas para sintetizar nuevos superconductores de materiales cerámicos avanzados, se puede razonar el efecto biológico de un catión u otro, o predecir qué elementos se forman por desintegración radiactiva a partir de los núcleos de un elemento concreto, por poner algunos ejemplos.



Fig. 11. Tabla periódica en chino simplificado.

Contribución española a la construcción de la tabla periódica.

La aportación de nuestro país en la construcción de la TP no es menor: sólo hay del orden de una decena de países cuyos científicos han descubierto 2 o más de los elementos. Uno de ellos es España, con tres, como se resalta en el segundo sello de la Fig. 1. En todo caso, a veces es difícil precisar quién descubre un elemento químico: ¿La primera persona que lo detecta? ¿El primero que lo aísla? ¿El que lo caracteriza en primer lugar? Incluso, puede haber controversias porque se realizó el descubrimiento en distintos sitios y los protagonistas tuvieron diferente fortuna en la difusión de su conocimiento. Tampoco todos los países actuales se corresponden con los de hace cien o doscientos años, ni parece conveniente hacer de esta cuestión un ejercicio de nacionalismo. Pero sí por cercanía y por honrar su labor, se recogen aquí,

brevemente, algunos aspectos de la participación española en la construcción de la TP.

El único elemento cuyo nombre y símbolo tienen etimología española es el platino, que descubrió el marino, aventurero y científico ilustrado Antonio de Ulloa y de la Torre (Sevilla, 1716 – Isla de León, Cádiz, 1795). Con solo 19 años participó, acompañado de Jorge Juan (Monforte del Cid, Alicante, 1713 – Madrid, 1773), en la *Misión Geodésica Hispano-francesa*, que transcurrió durante nueve años principalmente por la *Real Audiencia de Quito* (hoy Ecuador), para medir la longitud de un arco de meridiano y resolver así un problema de geodesia sobre la esfericidad de la Tierra. Allí, Ulloa se dio cuenta de que una impureza de oro y plata, que se denominaba “platina” por su similitud con la plata, era un nuevo metal [16]. El wolframio fue descubierto en el *Real Seminario Patriótico de Vergara* por los hermanos Elhuyar y Lubice: Juan José (Logroño, 1754 – Bogotá, 1796) y Fausto Fermín (Logroño, 1755 – Madrid, 1833). Con formación europea, ambos desarrollaron su carrera científica en América (Juan José en el Virreinato de Nueva Granada, actual Colombia, y Fausto en el de Nueva España, actual México). Aunque la RAE admite también la denominación de tungsteno, la comunidad de químicos españoles suele preferir que se designe wolframio. Andrés Manuel del Río Fernández (Madrid, 1764 – Ciudad de México, 1849) estudió en el Instituto (entonces *Real Colegio*) de San Isidro (ya mencionado), en la Universidad de Alcalá, en la Escuela de Minería de Almadén y en varios países europeos. Fue de profesor a México donde, aparte de importantes logros en mineralogía, descubrió el elemento que se conocería como vanadio.

Otras aportaciones españolas, más recientes, son las de Blas Cabrera y Felipe (1878-1945) en el estudio de las propiedades magnéticas de los átomos, las de Enrique Moles Ormella (1883-1953), que llegó a ser secretario de la *Comisión de Pesos Atómicos* de la IUPAC, y las de Miguel Antonio Catalán Sañudo (1894-1957), por su interpretación de la configuración electrónica y la elaboración de una TP basada en los espectros atómicos.

Dentro de la celebración de 2019 como *Año Internacional* de la TP, las actividades emprendidas en España han alcanzado un eco internacional. Por ejemplo, en la fachada de la Facultad de Química de la Universidad de Murcia se encuentra la considerada como la TP más grande del mundo y su fotografía se incluye en la información aportada por la IUPAC sobre la efeméride.

El futuro previsible de la tabla periódica.

Aparte de ciertas controversias pendientes ya citadas, a la hora de elegir la ubicación y disposición de algunos elementos en la TP, podemos afirmar que están descubiertos los que constituyen, al menos, la materia accesible a nosotros (no quedan “huecos” como en la propuesta inicial de Mendeléiev). Hay que destacar aquí que, a partir del uranio (de número atómico 92) los siguientes elementos (transuránidos) son sintéticos e inestables, con una vida media tan pequeña que no cabe pensar en la preparación de sustancias de interés cotidiano, aunque sí muy relevantes desde el punto de vista de la física nuclear (incluyendo aplicaciones médicas). Para los elementos de número

atómico 119 y 120, que aún no se han obtenido y que, provisionalmente, la IUPAC les asigna los símbolos de Uue (ununennium) y Ubn (unbinilio), por la forma de citar en latín sus números atómicos, respectivamente, es de esperar que sigan el procedimiento actual de asignación de nombre cuando se admita su autoría: la IUPAC da un tiempo para que los investigadores protagonistas del descubrimiento (hoy en día centros de investigación y no investigadores actuando de forma aislada) propongan un nombre relacionado con términos geográficos, nombres de científicos, un concepto o personaje mitológico (incluyendo objetos astronómicos), un mineral o sustancia similar, o una propiedad del elemento. Así se nombraron en 2016 los últimos elementos: nihonio (113), moscovio (115), teneso (117) y oganesón (118). Ojalá alguna vez se sugiera el nombre de algún científico español, por ejemplo los citados en este texto. El geólogo inglés George W. Featherstonhaugh, que desarrolló su labor en Estados Unidos, sugirió que el vanadio se nombrase “rionio” en honor a su descubridor, pero nunca se ha considerado la idea. También merecen reconocimiento, en cuanto a nomenclatura de elementos, personajes como Lavoisier, Dalton, Faraday y Moseley, entre otros, y sería positivo atender a otras zonas geográficas: no parece justo que los países con más recursos y, por tanto, con una investigación más avanzada y consolidada, acaparen los nombres de los elementos.

Además de la propia nomenclatura, cuando se inicie el descubrimiento del elemento de número atómico 121 y posteriores, nos encontraremos con el mismo problema que la ubicación de los lantánidos y actínidos en la TP. A partir de ese elemento, se inician previsiblemente las configuraciones electrónicas con 9 orbitales g (en los que puede haber 18 electrones y, consecuentemente, ese mismo número de elementos). Es probable que se aborde, en su día, incluyéndolos debajo de los actínidos, también fuera de la tabla general.

De cualquier forma, la creatividad y los avances científicos igual consiguen que, dentro de un tiempo, alguien tenga la genialidad de ordenar los elementos de otro modo, atendiendo siempre a su periodicidad, aprovechando quizá la versatilidad de las herramientas digitales.

A modo de conclusión.

La historia de la construcción de la TP es un ejemplo de cómo se desarrolla la ciencia, una obra colectiva e internacional. El resultado es apasionante; es difícil imaginar en cualquier área del saber otro esquema tan fundamentado y con tanta información útil recogida en una tabla que entra en una hoja. Este icono universal no sólo es historia; es clave para entender muchos avances y progresos en multitud de áreas, como la física nuclear, la ciencia de los materiales, la biomedicina, etc.

Dada la limitación por las características de este texto, no se ha profundizado en muchos aspectos. Para lectores con deseos de seguir introduciéndose en el tema se recomiendan, aparte de los textos recogidos en las referencias, dos libros: el de Oliver Sacks [17], donde explica el atractivo que le supuso la tabla periódica en su infancia, aunque profesionalmente se

dedicó, ya de adulto, a la neurología, y el de Theodore Gray, con fotografías espectaculares sobre los elementos químicos y sus aplicaciones realizadas por Nick Mann y él mismo, que se recogen también en una página Web [18].

Ojalá la celebración del *Año Internacional de la Tabla Periódica*, en el sesquicentenario de la propuesta inicial de Mendeléiev, sirva para que se conozca más, especialmente por niños y jóvenes, esta joya de la humanidad, no siempre suficientemente valorada. Además, es buena ocasión para conocer y destacar la labor de la IUPAC, especialmente en la definición de la nomenclatura química, con ocasión de su centenario.

Referencias:

- [1] International Union of Pure and Applied Chemistry (IUPAC). The international year of the Periodic Table: A common language for science. <https://www.iypt2019.org/>
- [2] Pinto G. 2007. A postage stamp about the periodic table. *Journal of Chemical Education* 84: 1919.
- [3] Pinto, G. 2018. Modelos atómicos y propiedades periódicas de los elementos químicos: Recursos para el aprendizaje. *Alambique, Didáctica de las Ciencias Experimentales* 93: 41-45.
- [4] Pinto Cañón, G.; Castro Acuña, C. M.; Martínez Urreaga, J. Ed. 2006. Química al alcance de todos. Madrid. Pearson Alhambra.
- [5] Pinto, G.; Martín, M.; Calvo Pascual, M. A.; de la Fuente, A. 2019. Año internacional de la tabla periódica (2019): Una oportunidad para abordar contextos de didáctica e historia de la física y la química. *Revista Española de Física*, en prensa.
- [6] Scerri, E. Ed. 2013. La tabla periódica: Una breve introducción. Madrid. Alianza Editorial.
- [7] Esteban Santos, S. Ed. 2009. La historia del sistema periódico. Madrid. Cuadernos de la UNED.
- [8] Moeller, T. Ed. 1959. 3ª edición. Química inorgánica. Barcelona. Reverté.
- [9] Román Polo, P. Ed. 2002. Mendeléiev: El profeta del orden químico. Tres Cantos (Madrid). Nivola Libros y Ediciones.
- [10] Soddy, F. 1913. Intra-atomic charge. *Nature* 92: 399-400.
- [11] Brunning, A. Compound Interest. <https://www.compoundchem.com/>
- [12] Mans, C. 2010. La tabla periódica del edificio histórico de la Universidad de Barcelona. *Química e Industria* 587: 36-40.
- [13] Royal Society of Chemistry (RSC). Periodic table. <http://www.rsc.org/periodic-table/>
- [14] Winter, M. Webelements: The periodic table on the WWW. <https://www.webelements.com/>
- [15] Román Polo, P. 2018. Los nombres de los elementos químicos desde el siglo XVIII hasta nuestros días. En Ruiz-Berdún, D. (Ed.). Ciencia y Técnica en la Universidad: Trabajos de historia de las ciencias y de las técnicas, Vol. I. Universidad de Alcalá: Servicio de Publicaciones, pp. 173-184.
- [16] Pinto G. 2017. Antonio de Ulloa and the discovery of platinum: An opportunity to connect science and history through a postage stamp. *Journal of Chemical Education* 94: 970-975.
- [17] Sacks, O. 2006. El tío Tungsteno: Recuerdos de un químico precoz. Madrid. Anagrama.
- [18] Gray, T. 2009. The elements: A visual exploration of every known atom in the universe. Nueva York. Black Dog & Leventhal Publishers. Con información complementaria (imágenes, vídeos, etc. de los elementos) en: <http://periodictable.com/theelements/pages.html>

Gabriel Pinto Cañón

Doctor en Ciencias Químicas (especialidad de Química Física) por la *Universidad Complutense de Madrid*.

Catedrático de Universidad (área de Ingeniería Química) en la *Universidad Politécnica de Madrid*.

Presidente del *Grupo Especializado de Didáctica e Historia*, común a las *Reales Sociedades Españolas de Física y de Química*.

Scientix Ambassador.

Capítulo 3

La creación de los números: muchos pasos para un gran fin.

Una popular frase dice: <<Solo con matemáticas la humanidad no hubiese podido ir a la Luna, pero sin matemáticas tampoco>>. Este artículo quiere mostrar como los números han sido un gran paso para el progreso humano.

Números de todo tipo acompañan nuestra vida. Nunca los números estuvieron tan presentes en el mundo como ahora ((1), (2), (4), (5), (8)) y nunca antes hubo las facilidades de cálculo que hoy tenemos a nuestra disposición. Por ello es curioso hacer un pequeño recorrido histórico para descubrir cómo, lentamente, los números fueron inventados por la humanidad para resolver problemas concretos ((9), (12), (17)). En este apartado tendremos ocasión de ver claramente que el principio “cada problema tiene los números que se merece” ha sido cierto desde siempre. Los números están en el corazón de las matemáticas, en el nacimiento de la propia disciplina.

La historia de los números se funde con la historia de la humanidad y va unida, es sus inicios, a la ancestral e imprescindible necesidad de *contar*. A lo largo de miles de años se fue manifestando esta imperiosa voluntad por *cuantificar* lo más elemental: las ovejas de un rebaño, el resultado de una caza o pesca, los días de un trayecto, la llegada de la primavera o de la época de lluvias. *Contar* para recordar, para distribuir, para ordenar e incluso para hacer previsiones. Estas cuentas se formalizaron en muescas en maderas o huesos, en colecciones de piedras, en el uso de los dedos, en sonidos y finalmente en símbolos diversos que dieron luz a los sistemas de numeración.

Si *contar* y *ordenar* fueron los primeros objetivos numéricos pronto los números fueron básicos para *medir* (longitudes, superficies, volúmenes, tiempo,...) y para poder *calcular*.

Los “números” fueron necesarios antes que las “letras”. Según se describe en (21), el primer alfabeto posiblemente fue ideado por el semita Khebeded (entre 1990 y 1790 a.C.) en los yacimientos de turquesas de Serabit el-Jadim en el Sinaí (15).

Números con muescas, piedras, dedos,...

Desde el Paleolítico se conservan huesos con muescas (16). Posiblemente también en maderas (hoy desaparecidas) se practicaron incisiones. Y en las paredes de las cuevas se hicieron marcas. El paso del nomadismo al cultivo

facilitó sin duda este avance. Las muescas dejaban constancia de un recuento (rebaño, comida, caza, enemigos,...) y por tanto acostumbraban a ser sucesiones de marcas. En algunos casos las muescas realizadas en trozos de cañas o maderas permitían, al ser cortado el soporte por la mitad, que *vendedor* y *comprador* pudieran guardar testimonio gráfico de la transacción o acuerdo numérico, asegurándose que en cualquier momento pudieran juntarse los dos trozos y verificar cual había sido el resultado pactado.



Fig.1. Hueso de Ishango.

El testimonio más espectacular de contabilidad avanzada con muescas en un hueso es el caso del llamado *hueso de Ishango* (posiblemente con más de 20.000 años de antigüedad) que fue hallado en las orillas del lago Eduardo en Zaire y hoy se encuentra en el Museo de Ciencias Naturales de Bruselas. Destacan en este hueso las

secuencias “agrupadas” de muescas, algunas de las cuales corresponden a *números primos*. En efecto, aparecen los primos entre 10 y 20; impares agrupados (9, 11, 13, 17, 19, 21), y algunas secuencias suman 60 y otras 48 (múltiplos de 12).

Otra forma ancestral de hacer cuentas y guardar el resultado fue tener una colección de piedras (unidades) y guardar las cantidades en bolsitas o cuencos. El pastoreo usó siempre este sistema para verificar si todo el ganado estaba allí o se había perdido algún animal.

Pero el comercio, las medidas agrícolas, etc. obligaron a perfeccionar el sistema. Lo primero fue tomar piedras pequeñas como unidades, piedras mayores para decenas, otras mayores para centenas, etc. También con arcilla se hicieron símbolos: bastoncillos o pequeños conos para unidades, bolitas para decenas, discos o conos grandes para centenas, esferas grandes para millares,... En estas formas de representar cantidades, las piezas se llamaron en latín *calculi* y de ahí derivaría nuestro concepto actual de *cálculo*.

Sumerios y elamitas perfeccionaron aún más este sistema. En concreto los sumerios (en base 60) usaron formas de conos y esferas pero con el criterio de que si se perforaban estas figuras

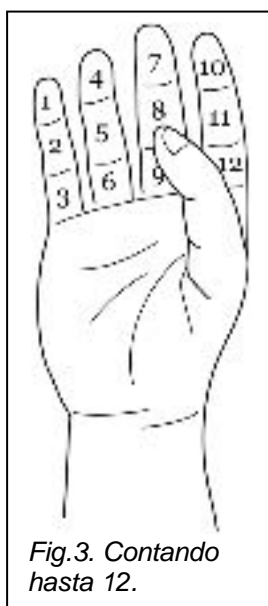


Fig.2. Piezas en arcilla con valores numéricos.

con un agujero el valor se multiplicaba por 10. Repitiendo estas piezas se componían los números. Los elamitas usaron un palito para 1, una bolita para 10, un disco para 100, un cono para 300 (60x5) y un cono mayor perforado para 3000 (60x5x10).

En bolas huecas se guardaban los números compuestos con los *calculi* cuando interesaba dejar constancia de un acuerdo, una posesión, un débito, etc. Y a partir de aquí surgió la genial idea de evitar el contar el interior de estas

bolas recipientes, marcando simbólicamente sobre ellas lo que estaba en el interior.



Nuestros más remotos antepasados también hicieron su revolución “digital” al darse cuenta de que los dedos eran un instrumento valioso para contar (6). Con 10 dedos en las manos, la base 10 estaba servida, pero curiosamente con una sola mano se podía contar hasta 12.

Al considerar los 10 dedos de las manos y los 10 de los pies también motivó a ciertas culturas americanas al uso de la base 20.

Números y palabras.

En paralelo a muescas, piedras, nudos, dedos y primeros símbolos, también surgieron sonidos y palabras para designar números. Mucho antes de que los símbolos llevaran a la numeración simbólica escrita hubo numeración oral. Los sonidos numéricos fueron en origen una aproximación más cualitativa que cuantitativa, más pensada para distinguir pequeñas cantidades (uno, dos, muchos) que para precisar cifras grandes.

El diccionario español de la Real Academia Española (RAE) es una trabajada lista de 93.111 palabras (con 19000 americanismos) ordenadas alfabéticamente y con precisas definiciones. Las palabras de un idioma son el código de comunicación básica para hablar y escribir. Lo único que debemos constatar aquí es que el diccionario también es una lista de palabras con números. Hay infinitos números y por tanto los artífices del diccionario han tenido que hacer una selección de nombres de números básicos en la numeración:

- Los *numerales cardinales* expresan las cantidades discretas, desde el cero, uno, dos,....., millón, millardo, billón, trillón, cuatrillón, quintillón.
- Los *numerales ordinales* expresan el orden en secuencias ordenadas: primero, segundo,....., millonésimo.
- Los *numerales multiplicativos* permiten nombrar multiplicaciones: doble, duplo, triple,....., céntuplo.
- Los *numerales fraccionarios o partitivos* permiten expresar partes o divisiones. En substantivos: mitad o medio, tercio,...., dosmillonésimo o dosmillonésima; o en adjetivos: medio, media, tercera (parte),...., dosmillonésima (parte).

Unas joyas numéricas históricas.

Por tierras de Babilonia floreció durante siglos la denominada cultura babilónica a la cual contribuyeron también multitud de pueblos (sumerios, acadios,

caldeos, asirios,...) durante un largo periodo, desde el 2.100 a. C. hasta el 200 a. C. (19).

Con un estilete largo, cuya sección final era un triángulo isósceles, los babilonios crearon sus símbolos numéricos al combinar diversas presiones del estilete sobre tablillas de barro (*caracteres cuneiformes*) que luego en un horno dejaban el contenido aritmético listo para la posteridad. Para números menores que 60 se usaban solo dos símbolos, para 1 y 10, y un signo operativo.

Así pues entre 1 y 60 el sistema era aditivo con valores basados en la base 10. Lo sorprendente es que sumerios y babilonios usaran el sistema de base 60 posicional para números mayores (si bien no tenían símbolo para cero, lo que creaba ciertas ambigüedades). Por ejemplo: para el 201 al ser $201=3 \times 60 + 21$ se especificaba el símbolo de 3 seguido del de 21.

Este sistema permitió resolver todo tipo de anotaciones comerciales, intereses, censos, medidas, etc., e incluso iniciar estudios geométricos y algebraicos. El cálculo babilónico de que el año tenía 360 días seguramente influyó en esta adopción de la base 60. A ello se debe también la división del círculo en 360 partes. La idea de usar la base 12 para medidas de todo tipo duró muchos siglos y fue común a muchas culturas, siendo de hecho el sistema métrico decimal el principal introductor del 10 para crear múltiplos y submúltiplos en medidas. El hecho de que $12 = 2 \times 6 = 3 \times 4$ facilitaba que mitades, terceras partes, cuartas partes, etc. fueran fáciles de determinar. No deja de ser curioso que durante siglos el uso del 10 fuera común en aritmética a la vez que el 12 tuviese un protagonismo especial en medidas.

Hoy en día no solo se venden huevos por docenas sino que las circunferencias se siguen dividiendo en 360° y las medidas temporales siguen fieles al sistema babilónico: el reloj es dividido en 12 horas, cada hora en 60 minutos y cada minuto en 60 segundos... y los años siguen con 12 meses. La Sociedad Duodecimal americana reivindica aún hoy la base 12 como referente mundial.

Entre las trescientas tablillas matemáticas babilónicas destaca como joya de la corona una catalogada (en la americana Universidad de Columbia) como Plimpton 322 en honor a su donante G.A. Plimpton. Datada entre el 1900 a. C. y el 1600 a. C. y descifrada por Neugebauer y Sachs en 1945, la tablilla en cuestión contiene una extraordinaria sorpresa.



Fig.4. La tablilla Plimpton 322.

Aparecen los valores de los lados enteros de triángulos rectángulos. Una columna da catetos, otra las hipotenusas y la otra los catetos que acaban de completar dichos triángulos rectángulos. Así en la primera fila si la hipotenusa es 169 y un cateto es 119 resulta el otro 120 dado que $169^2 = 119^2 + 120^2$.

La sorpresa fue enorme. De una, aparentemente intrascendente lista de números afloraron unos resultados de cálculo avanzado (que hoy

denominaríamos de *tripletes pitagóricos*) mostrando que sabían generar dichos números y que por tanto mucho antes de Pitágoras ya se conocía esta propiedad. Hoy sabemos que para generar números enteros a, b, c tales que $a^2+b^2 = c^2$ se pueden usar números enteros u, v primos entre sí con $u > v$ y tomar $a = 2uv, b = u^2-v^2$ y $c = u^2+v^2$. Esto es exactamente lo que, en casos concretos, hace esta sorprendente tablilla.

¿Qué llevó a los autores de este listado a interesarse por este tema? Todo un misterio que da pie a numerosas conjeturas: ¿era un material didáctico de ayuda para educadores que quisieran proponer problemas a distinguidos alumnos?, ¿era una investigación numérica especulativa?, ¿tenía relación con los triángulos rectángulos subyacentes a formas perpendiculares en arquitectura?,....

Más allá de los tripletes pitagóricos de la Plimpton 322, en muchas otras tablillas babilónicas se da un desarrollo avanzado de cálculos algebraicos y se resuelven algunas ecuaciones de grado dos, tres y cuatro. Estos cálculos les permitieron calcular aproximaciones de determinados valores como:

$$\sqrt{2} \approx (17/12)(1/\sqrt{2}) \approx 17/24$$

o la espectacular aproximación de $\sqrt{2}$ con fracciones de denominador potencias de 60:

$$\sqrt{2} \approx 1+(24/60)+(51/60^2)+(10/60^3) \approx 1,4142155$$

Los papiros egipcios matemáticos.

Hacia el 3000 a.C. los egipcios empezaron a desarrollar un arcaico sistema de numeración basado en la base diez y en la repetición simple (de derecha a izquierda) de las representaciones simbólicas que dieron al 1, 10, 100, 1000, etc. En la etapa jeroglífica estas representaciones se correspondían con dibujos de determinados elementos: el uno era un trazo vertical, la decena una herradura (U invertida), la centena una espiral enrollada, el millar una flor de loto, la decena de mil un dedo levantado, la centena de mil una rana o renacuajo, el millón un hombre arrodillado con los brazos levantados al cielo (¡símbolo de gran actualidad!). Con este tipo de numeración solo las sumas y restas eran fáciles de realizar, así como las limitadas multiplicaciones o divisiones por dos o por diez.

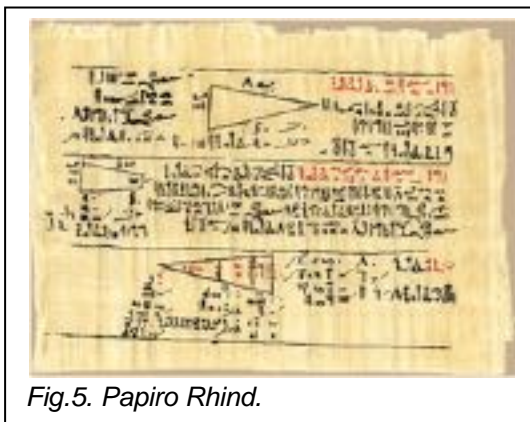


Fig.5. Papiro Rhind.

En una mina de Tebas apareció en el siglo XIX un papiro de 6 m de largo por 33 cm de ancho que hacia el 1650 a.C. había sido elaborado por el escriba Ahmes. Fue comprado en Luxor en 1858 por el egiptólogo escocés A. Henry Rhind y posteriormente adquirido por el Museo Británico de Londres que actualmente lo exhibe con gran satisfacción junto a su amplio botín

arqueológico acumulado. No deja de ser curioso que en lugar del nombre Ahmes del escriba, el papiro mantenga el nombre de su comprador (¡es excepcional que los ingleses pagasen algo para quedarse con ello!).

Amhes empieza el papiro con un titular digno del mejor marketing científico:

<<Cálculo exacto para entrar en conocimiento de todas las cosas existentes y de todos los oscuros secretos y misterios>>.

Amhes recopiló una lista de 87 problemas y sus soluciones, posiblemente para facilitar el aprendizaje matemático de la contabilidad y la medición. No incluye métodos ni fórmulas, sino soluciones a problemas cotidianos muy concretos: repartos de varias barras entre 10 hombres, operaciones con fracciones, ecuaciones de primer grado, progresiones numéricas, volúmenes de grano, áreas de terrenos, pesos, casos de triángulos, rectángulos, círculos y pirámides, divisiones y repartos proporcionales, etc.

Curiosamente, para facilitar los cálculos, el papiro incluye una tabla sobre como descomponer fracciones del tipo $2/n$ (desde $n = 5$ hasta $n = 101$) en sumas de fracciones de numerador 1, por ejemplo:

$$2/7 = 1/4 + 1/28, \quad 2/97 = 1/56 + 1/679 + 1/776.$$

Siempre sorprenden las grandes pirámides egipcias y los templos construidos. La grandeza de sus medidas y la notable posición de las inclinaciones y la localización geométrico-solar de estas edificaciones demuestran un alto nivel de pericia en las medidas y en su implementación técnica.

Pero la “gran pirámide” es una fórmula que se halla en el llamado *papiro de Moscú*. El papiro de Moscú adquirido por Golenishechev ha acabado siendo denominado con el nombre de la ciudad que desde 1917 lo acoge en su museo de Bellas Artes. Pieza de 5 m de largo por tan solo 8 cm de ancho, consta de 25 problemas y sus soluciones.

La aportación más interesante de esta joya (escrita con caracteres hieráticos hacia el 1890 a. C.) es la expresión exacta para el volumen de un tronco de pirámide de bases cuadradas, lo que llevó a E.T. Bell a calificar este cálculo como “la mayor pirámide de Egipto”. Concretamente, para un tronco de pirámide de altura 6 y lados de las bases cuadradas 4 y 2 el escriba dice:

*Elevar al cuadrado 2 y 4
Multiplicar 2 por 4
Sumar los resultados anteriores
Multiplicar el resultado anterior por un tercio de 6
El resultado es 56
Ves, el 56, lo has calculado correctamente.*

todo ello se corresponde con la fórmula exacta:

$$V = (a^2 + b^2 + ab)h / 3$$

Números escritos en base 10.

Al final del proceso surgieron los símbolos escritos numéricos. Los símbolos matemáticos son en sí mismos identificadores de la propia disciplina constituyendo lo que hoy, casi podríamos denominar, una imagen corporativa de la reina de las ciencias.

La primera virtud simbólica es sin duda la *simplificación*: es más rápido y simple usar determinados símbolos que palabras, dibujos, frases, etc. Pero la virtud más trascendente de los símbolos es la posibilidad *operativa* que a menudo nos facilitan. La diferencia entre la frase “el doble de mi edad es ciento catorce” y su representación “ $2x = 114$ ” es que la descripción simbólica da pie a desenmascarar rápidamente la incógnita: $x = 114/2 = 57$.

Los símbolos matemáticos más ancestrales fueron, por supuesto, los numéricos. Desde las primitivas representaciones que acabamos de ver al actual sistema de escribir los números naturales han transcurrido muchos miles de años que han visto la evolución de las cifras (16). Detrás de nuestra lista de queridos numerales

0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12,...

se esconden las muy diferentes formas de anotar simbólicamente los números naturales a lo largo de varios milenios. De todas estas representaciones de los números sobreviven hoy tres: la romana, la hindú que es la nuestra y la árabe.

Los romanos idearon un sistema numérico de base 10:

I, II, III, IV, V, VI, VII, VIII, IX, X,...

que era útil para expresar cantidades moderadas pero dado su carácter aditivo no fueron adecuados para realizar operaciones como la multiplicación o la división.

Los hindús tuvieron el ingenio de desarrollar un sistema de base 10 con cero 0 y con escritura posicional. Así el 3405 era 3 veces el 1000, 4 veces el 100, no había decenas y 5 unidades. Con este sistema las operaciones aritméticas pudieron facilitarse. Fueron los árabes los que se incorporaron este sistema, realizando meritorios hallazgos algebraicos y transmitiendo dicho sistema a Occidente. Ya en el siglo XIII fue Leonardo de Pisa, Fibonacci, quien con su *Libro de los ábacos* promovió el conocimiento del sistema hindú. Pero también a través de España dicho sistema tuvo difusión europea.

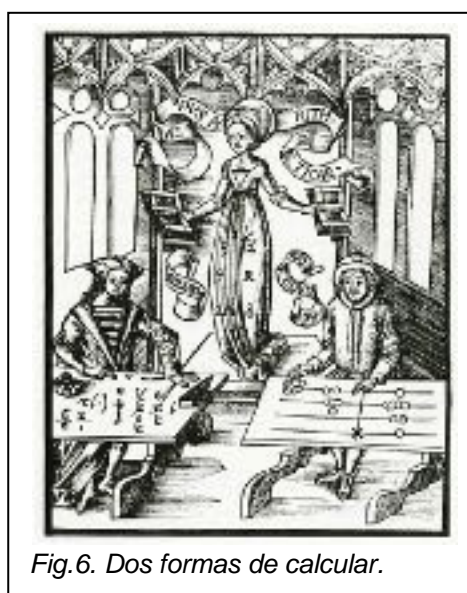


Fig.6. Dos formas de calcular.

En el grabado puede verse el

contraste entre el personaje de la derecha haciendo cálculos con *un ábaco* de piedras y el de la izquierda usando números hindús.

Los cálculos aritméticos se facilitaron durante años con el uso de los ábacos. Posteriormente con los algoritmos árabes manuales (los nuestros) y otros ingenios como las “*reglas de cálculo*” y las *tablas de logaritmos*. Ya en 1620, William Gunter desarrolló un sistema de reglas con números logarítmicos para facilitar el cálculo manualmente. En 1622 William Oughtred inventó unos círculos giratorios y finalmente en 1633 dio a luz la idea de unas reglas (rectas) para calcular. Estas reglas eran aún un instrumento efectivo de cálculos en 1970. Murieron con las calculadoras.

Números de muchos tipos.

La creación de los *números naturales* $N = \{0, 1, 2, 3, 4, \dots\}$ ya fue un gran avance que duró muchos siglos. Con los números naturales los temas de contar y hacer los cálculos básicos de la Aritmética ya estaban resueltos. Pero con el progreso de la humanidad pronto surgieron nuevos retos que exigieron la creación de otras clases de números. Vale la pena consultar en internet la web (<https://oeis.org>) de la *On-line Encyclopedia of Integer Sequences*, una enciclopedia enorme de las sucesiones de enteros.

Los *números enteros* $Z = \{\dots, -5, -4, -3, -2, -1, 0, +1, +2, +3, +4, +5, \dots\}$ amplían los naturales (que se identifican con los enteros positivos) y añaden los *números negativos*, haciendo posible las restas, los cálculos mercantiles positivos y negativos y hoy el uso de escalas con positivos y negativos a la vez: grados de temperatura, numeración de ascensores, coordenadas simétricas. El matemático indio Bramagupta, en el 628 de nuestra era, considera las dos raíces de las ecuaciones de segundo grado, aunque una de ellas sea negativa o irracional. En su obra aparecen los números positivos, negativos y el cero, que él llamaba *los bienes, las deudas y la nada*.

Los *números racionales o fraccionarios* $Q = \{m/n \text{ con } m, n \text{ enteros, } n \neq 0\}$, que ya habían aparecido en Egipto, al relacionar dos enteros permiten establecer proporciones (regla de tres), hacer muchos cálculos, resolver ecuaciones del tipo $nx+m = 0$, invertir enteros ($1/n$), dar valores intermedios entre enteros como calcular medias aritméticas $(n+m) / 2$, etc. A dichos números fraccionarios m / n por división se les puede asociar una representación con decimales ($1/3 = 0,3333\dots, 1/4 = 0,25, \dots$) con la peculiaridad que a partir de algún decimal aparece siempre un grupo de cifras que se va repitiendo indefinidamente.

Los números *irracionales* son los que conjuntamente con los racionales completan los *números reales* R que se corresponden con los puntos de una recta (3). Entre los irracionales destacan los *algebraicos* como las raíces cuadradas, cúbicas y de cualquier orden $\sqrt{2}, \sqrt{3}, \dots$ (13) o sus combinaciones con fracciones como es el caso del popular *número de oro (divina proporción)* ((11), (14), (18), (20)):

$$(1+\sqrt{5})/2 \cong 1,618$$

Otros irracionales denominados *trascendentes* tienen enorme importancia por su interés matemático. Destacan por su importancia histórica el *número pi* π (7),

$$\pi = 3,1415926535897932384626433\dots$$

razón entre el perímetro de una circunferencia y su diámetro, y el *número e* ligado al cálculo de logaritmos y a series numéricas, siendo e límite de la sucesión numérica $(1 + 1/n)^n$.

Los irracionales tiene también su representación mediante decimales pero en su desarrollo nunca aparece un grupo repetido de cifras. Por curiosidad incluimos los primeros 1000 decimales del número de oro

```
1.61803398874989484820458683436563811772030917980576286213544862270526046281890
244970720720418939113748475408807538689175212663386222353693179318006076672635
443338908659593958290563832266131992829026788067520876689250171169620703222104
321626954862629631361443814975870122034080588795445474924618569536486444924104
432077134494704956584678850987433944221254487706647809158846074998871240076521
705751797883416625624940758906970400028121042762177111777805315317141011704666
599146697987317613560067087480710131795236894275219484353056783002287856997829
778347845878228911097625003026961561700250464338243776486102838312683303724292
675263116533924731671112115881863851331620384005222165791286675294654906811317
159934323597349498509040947621322298101726107059611645629909816290555208524790
352406020172799747175342777592778625619432082750513121815628551222480939471234
145170223735805772786160086883829523045926478780178899219902707769038953219681
9861514378031499741106926088674296226757560523172777520353613936
```

Lograda la construcción del cuerpo numérico de los números reales R aparecen aún problemas no resolubles con dichos números como es el caso de las raíces cuadradas de enteros negativos ($\sqrt{-2}$, $\sqrt{-3}$,...) lo que hace que ecuaciones simples como $x^2 + 1 = 0$ no tengan raíces reales. Ello llevó a ampliar los números reales y definir el cuerpo de los *números complejos*

$$C = \{a + b i, a, b \text{ reales y } i = \sqrt{-1}\}$$

Muchos otros conjuntos numéricos como los *cuaterniones* se han ido construyendo en el los últimos dos siglos.

Cuentas binarias en base 2 y computación.

¿Qué tienen en común: un ordenador, un fax, una televisión digital, Internet, un móvil, una foto digital,...?, todo el secreto es la base 2, el *sistema binario* de escribir números usando sólo 0's y 1's como factores de las potencias de 2:

$$1 = 1; 2 = 1 \times 2 + 0 \times 1 = 10; 3 = 11; 4 = 100; \dots$$

Hoy es posible representar voces, colores, imágenes, letras,... con grandes cantidades de ceros y unos y esto explica cómo se puede "transmitir" cualquier cosa, de la voz a un escrito, pasando por los hilos o a través del aire. ¡Los ordenadores adoran el sistema binario para operar y almacenar información!

Un *bit* (viene del inglés *binarydigit*) es la unidad de información más pequeña: es dar un 0 o un 1; con 4 bits se tiene un *nibble* o cuarteto y con 8 bits se tiene un *byte*...

Vale la pena notar que el sistema binario no nació con los ordenadores sino mucho, mucho antes. Se han encontrado ejemplos indios y chinos de hace más de 2000 años donde expresaban números usando potencias de 2. En el siglo XVII la cosa va a más, Francis Bacon (1561-1626) habla de expresar (hoy llamaríamos *codificar*) palabras con números binarios y el gran matemático Gottfried Wilhelm Leibniz (1646-1716) escribe la conocida obra "Explication de l'Arithmétique Binaire" donde ya usa el 0 y 1 para expresar números en base 2. A mitad del siglo XIX George Boole se interesó en temas de lógica, con los valores de verdad o falsedad de las proposiciones (1 y 0) creando la llamada *álgebra de Boole*. Quería, en cierto modo, ligar la lógica con un sistema formal de cálculo binario. Poco podía sospechar el impacto que esto tendría años después.

Claude Shannon (1916-2001) que era ingeniero eléctrico fundador de la *teoría de la información* y George Stibitz (1904-1995), considerado padre de los ordenadores digitales, implementan en 1937 las ideas teóricas booleanas en la creación de circuitos, en las primeras computadoras que operaban usando el sistema binario. John von Neumann (1903-1957), John Mauchly (1907-1980) y Norbert Wiener (1894-1964) hacen ya posible una novedosa "cibernética". Y de entonces a hoy llegamos a una computación que ha cambiado el mundo. La máquina de vapor hizo posible la revolución industrial; los ordenadores hacen ahora posible la revolución del conocimiento.

Una figura clave en las ciencias de la computación fue el británico Alan Turing (1912-1954) abriendo con gran creatividad nuevos horizontes a las *ciencias de la computación*. Durante la Segunda Guerra Mundial dirigió una división de la Inteligencia británica para descifrar (con nuevas máquinas e ideas sobre códigos) los mensajes secretos de la marina alemana. Entre 1945 y 1952 desarrolló aportaciones muy originales sobre programación, redes neuronales e inteligencia artificial. Las ideas de Turing se avanzaron a la tecnología de su tiempo pero contribuyeron enormemente al desarrollo posterior de las máquinas computacionales y a la reflexión sobre sus límites y posibilidades.

Como dice Brian W. Kernighan, detrás del mundo computacional hay tres grandes ideas (máquinas, programas y comunicación) y un principio clave que es el de la digitalización, la reducción de todo tipo de "informaciones" a números binarios con ceros y unos.

Gran repercusión tuvo un artículo de Burks, Goldstein y von Neumann de 1946 donde aseguran poder "trabajar" (introducir, manipular, almacenar, enviar,...) solamente con números. Si bien la "estructura lógica" del ordenador se ha mantenido, a nivel tecnológico la evolución en formas, tamaños, potencia y precios, ha sido sorprendente desde los años cincuenta. Desde las pantallas de fósforo y los grandes tubos a los actuales circuitos integrados, con millones de transistores en un "chip", y los procesadores múltiples, la evolución ha sido espectacular siendo hoy posible hacer, en un mini portátil o un móvil, operaciones que no se podían soñar con máquinas que ocupaban una planta.

Según la conjetura de Moore, cada uno o dos años se podría doblar la capacidad de los computadores y así ha sido.

Hoy sigue avanzando el conocimiento de millones de decimales de pi, poniendo con ello a prueba las capacidades computacionales del momento. Mirar atrás y visitar algunas aproximaciones históricas de π nos permite revivir algunos momentos en donde conocimiento matemático e ingenio calculador se fundieron. El número pi es un fiel testimonio que ha calibrado el desarrollo matemático a lo largo de la historia (7). A partir de 1947 las calculadoras y ordenadores abrieron nuevas perspectivas a la pi-manía.

El imperio de los códigos.

Un código es cualquier grupo de símbolos que represente una información (10). Entendido así los sistemas de numeración, las palabras o frases de los idiomas, los jeroglíficos egipcios, etc. eran ya códigos. Pero el avance científico permitió que nuevos códigos fueran desarrollándose: el código Morse, el código Braille, el código genético del ADN, el código musical para escribir partituras, etc. a la vez que nuevos medios para comunicar estos códigos fueron naciendo (telégrafos, teléfonos, instrumentos musicales, radio, televisión, computadoras,...).

En el siglo XX los códigos han adquirido una importancia estratégica al formar parte esencial de la información característica de este siglo, desde CD musicales a lejanas sondas espaciales que envían datos, desde módems a móviles, desde Internet a la televisión de alta resolución. Y detrás de estos códigos hay, esencialmente, números.

En los productos comerciales, en el ISBN de los libros, en las etiquetas de almacenamiento, en los identificadores de los exámenes de selectividad, en cheques bancarios, en tarjetas, en envíos postales,... por todas partes aparecen códigos de barras: una información numérica precisa e identificadora es convertida a una secuencia de barras blancas y negras que pueden ser leídas de forma óptica y, recuperando la información, esta puede almacenarse o manipularse. En estos códigos aparecen dígitos de control asociados a los dígitos informativos y que permiten a los lectores ópticos detectar posibles errores.

La criptografía y los números primos.

Crear y romper secretos ha sido siempre un tema de interés. En *criptografía* se trata de inventar y romper códigos secretos. Curiosamente en la criptografía actual juegan un papel crucial los *números primos*, naturales solo divisibles por ellos mismos y la unidad (1, 2, 3, 5, 7,...).

Nótese que dado un número n puede ser muy complicado determinar si n es primo o no, pues debemos garantizar que no es exactamente divisible por ningún número menor que él. Y ello puede ser una labor muy, muy fatigosa.

También es tedioso dar un número (por ejemplo 3024) y descomponerlo en dos factores ($3024 = 36 \times 84$).

La sorpresa es que gran parte de las sofisticadas técnicas, para garantizar hoy en Internet la seguridad en la transmisión de datos confidenciales, se fundamentan en esta dificultad (humana y computacional) para descomponer números. Si se disponen de dos primos grandes p y q y se multiplican (¡fácil!) se obtiene un número elevado $m = p \cdot q$. Para el que quiera descomponer m el trabajo será largo si no conoce ni p ni q , pues al ser ambos primos, m no tendrá otros divisores. Este tipo de ideas han dado lugar a la denominada Criptografía de clave pública (y métodos como los de R. Merkle, W. Diffie y M. Hellman; el RSA de R. Rivest, A. Shamir y L. Adleman; el método Elgamal; etc.).

Si en el momento de enviar los dígitos de una tarjeta de crédito una parte de ellos queda visible a posibles piratas informáticos pero otra parte D de los dígitos se transmite multiplicada por un número primo grande P conocido por el receptor, cuando éste reciba el resultado $D \cdot P$ dividirá por P y tendrá D al instante. Si el pirata solo puede “ver $D \cdot P$ ” tendrá trabajo para hallar D pues además no sabe nada ni de D ni de P .

Conclusión.

En este breve recorrido histórico por los números se aprecia como el sistema numérico ha sido el resultado de un largo proceso. El resultado es hoy magnífico por lo que podemos afirmar que crear los números fue realmente un gran paso para la humanidad.

Actualmente la *Teoría de números* es una activa y brillante rama de la matemática que sigue investigando temas numéricos de gran calado. Pero además las aportaciones de los números al progreso tecnológico y científico están en la base de la actual revolución del conocimiento. Grandes pasos futuros serán posibles gracias a los números. Afortunadamente, lo mejor está aún por venir.

Bibliografía:

- (1) Alsina, C., 2008, *Vitaminas Matemáticas. Cien claves sorprendentes para introducirse en el fascinante mundo de los números*, Barcelona, Editorial Ariel.
- (2) Alsina, C., 2008, *Textos del Catálogo ¡Números! Utilidad y magia en la vida cotidiana*. Barcelona, Cosmocaixa-Fundació "La Caixa".
- (3) Alsina, C., 2009, *El teorema de Pitágoras. La aparición de los irracionales*, Barcelona, RBA Coleccionables.
- (4) Alsina, C. (col. Cristina Macía), 2016, *Mateschef. Un Sofrito de Números y Formas para Chefs y Gourmets*, Editorial Ariel.
- (5) Alsina, C., 2017, *Todo está en los números*, Barcelona, Editorial Ariel.
- (6) Antoñana, H., 1985, *La danza de los números*. Bilbao, DDB.
- (7) Beckmann, P., 1971, *A History of Pi*, New York, The Golem Press.
- (8) Bolt, B., 1991, *Mathematics meets technology*, Cambridge, Cambridge University Press.
- (9) Boyer, C., 1999, *Historia de la Matemática*, Madrid, Alianza Editorial.
- (10) Churchouse, R., 2002, *Codes and Ciphers*, Cambridge, Cambridge University Press.
- (11) Corbalán, F., 2010, *La proporción áurea*. Barcelona, RBA Coleccionables.
- (12) Davis, P.J.&Hersh, R., 1988, *Experiencia Matemática*, Barcelona, Labor-MEC.
- (13) Flannery, D., 2006, *The Square Root of 2. A dialogue Concerning a Number and a Sequence*, New York, Springer.
- (14) Ghyka, M.C. 1977, *Estética de las proporciones en la Naturaleza y en las Artes*, Barcelona, Ed. Poseidón.
- (15) Goldwasser, O., 2010, *How the Alphabet Was Born from Hieroglyphs*, *Biblical Archaeology Review*, 36, núm.2, 40-53.
- (16) Ifrash, G., 1992, *Historia universal de las cifras*, Barcelona, Espasa.
- (17) Kasner, E.&Newman, 1972, J., *Matemáticas e imaginación*, México, Ed. Continental.
- (18) Livio, M., 2009, *La Proporción Áurea. La historia de phi, el número más sorprendente del mundo*, Barcelona, Editorial Ariel.
- (19) Neugebauer, O., 1957, *The Exact Sciences in Antiquity*, New York, Dover.
- (20) Pedoe, D., 1982, *La Geometría en el Arte*, Barcelona, Ed. Gustavo Gili.
- (21) Wagensberg, J., 2017, *Teoría de la Creatividad. Eclósión, gloria y miseria de las ideas*, Barcelona, Tusquets Editores.

Claudi Alsina

Doctor en Matemáticas.

Catedrático de matemáticas (jubilado) de la Universidad Politécnica de Catalunya.

Capítulo 4

Rayos X: hallazgo y consecuencias de una luz prodigiosa.

Con la inconfundible precisión científica de la lengua alemana, comenzaba Wilhelm Conrad Röntgen su presentación ante la *Sociedad de Física y Medicina de Würzburg* (Alemania) sobre el descubrimiento que fortuitamente protagonizó el 8 de noviembre de 1895:

Cuando se deja pasar la descarga eléctrica procedente de una bobina Ruhmkorff a través de un tubo de vacío de tipo Hittorf, o de un Lenard suficientemente evacuado, Crookes o cualquier otro tubo parecido, cubierto con una camisa ceñida de cartón negro y fino, y en la habitación totalmente a oscuras se deja una placa de papel previamente recubierta de una capa de cianuro de bario y platino, ésta se ilumina cada vez que se produce una descarga, independientemente de que la superficie recubierta esté apuntando o no hacia el tubo. Dicha fluorescencia se produce hasta 2 metros de distancia del aparato. Es fácil convencerse de que la fluorescencia proviene del aparato de descarga eléctrica y no de cualquier otra parte de los equipos.



Fig.1. Carátula de la presentación oficial en 1896 del descubrimiento de Röntgen ante la Sociedad de Física y Medicina de Würzburg (Alemania).

Con dicha descripción, publicada en 1896, Röntgen introducía ante la mencionada sociedad científica las circunstancias del hallazgo de lo que él mismo describió como **un nuevo tipo de rayos** (*Eine neue Art von Strahlen* [1], Fig. 1) totalmente desconocidos, de ahí que se les bautizara con la letra de las incógnitas (X), pero que, nada más y nada menos, posteriormente abrirían puertas inimaginables para la ciencia del momento.

El descubridor.

Wilhelm Conrad Röntgen, nacido el 27 de marzo de 1845, fue el único hijo de un fabricante de tejidos asentado en Lennep, cerca de Remscheid (Alemania). A sus tres años, la familia paterna emigró a los Países Bajos, estableciéndose primero en Appeldoorn hasta 1862, y luego en Utrecht, en donde al joven Wilhelm Conrad no se le permitió acabar

con su graduación escolar, ya que al parecer fue confundido con el autor de una caricatura sobre su maestro de clase, sin que quisiera revelar al culpable.

En 1865 Röntgen abandonó los Países Bajos para estudiar ingeniería mecánica en Zürich (Suiza), en donde se enamoró de Anna Bertha Ludwig, sobrina del escritor Otto Ludwig (1813-1865), y con la que se comprometió tras superar brillantemente su licenciatura. En 1860, después de obtener su doctorado en física, Röntgen acompañó como asistente a su profesor, el físico experimental August Kundt (1839-1894), primero a Würzburg y más tarde a Estrasburgo. En 1872, Wilhelm Conrad y Anna Bertha se casaron en Appeldoorn. Su feliz matrimonio, de casi cincuenta años de duración, permaneció sin hijos, de ahí que adoptaran a una sobrina de Anna Bertha.

Tras nueve años en Giessen como profesor ayudante, Röntgen volvió a Würzburg, en donde, manteniéndose fundamentalmente en la línea de las investigaciones de los físicos Heinrich Hertz (1857-1894) y Philipp Lenard (1862-1947), experimentó con los fenómenos que ocurren durante las descargas eléctricas en el seno de gases que se mantienen bajo una presión extremadamente baja en el interior de una ampolla de vidrio que contiene dos polos eléctricos. Cuando se aplica un alto voltaje entre los dos polos se observa una descarga en forma de haz luminoso formado por electrones, explicación desconocida en aquellas fechas y de ahí que se hablara simplemente de “rayos catódicos”. Las propiedades de estas descargas eran bastante bien conocidas cuando Röntgen comenzó sus experimentos, pero nunca imaginó que aquellos experimentos le llevarían al descubrimiento de una nueva y prodigiosa radiación.

A la edad de 50 años, concretamente el 8 de noviembre de 1895, Röntgen quedó perplejo ante una inexplicable observación ocurrida durante un experimento en su laboratorio de la Universidad de Würzburg. A unos pocos metros de la ampolla de vidrio con la que observaba las mencionadas descargas eléctricas había un papel impregnado con una sustancia especial. Cuando se generaban los rayos catódicos el papel comenzaba a brillar; pero lo que verdaderamente intrigó a Röntgen era que el papel todavía brillaba cuando tapaba totalmente la ampolla de descarga en el interior de un grueso cartón negro, de tal modo que el fogonazo de la descarga eléctrica no pudiera iluminar el papel impregnado. Tras replantearse detenidamente el resultado de las observaciones



*Fig.2. ¡Un resultado inesperado!
Descubrimiento de los rayos X en 1895.
Ilustración de Alejandro Martínez de Andrés,
CSIC, 2014.*

(Fig. 2) llegó a la conclusión de que el papel impregnado estaba siendo iluminado por un tipo de rayos totalmente desconocidos hasta aquel momento y capaces de traspasar el cartón, de ahí que los denominara con la letra de las incógnitas: rayos X. Guardó total silencio sobre su increíble hallazgo, hasta tal punto que comentó a su esposa que “si la gente supiera en lo que estoy

trabajando, pensaría que me he vuelto loco”. Pasó varias semanas casi sin dormir estudiando las propiedades de los rayos recién descubiertos, actuando con precisión y gran cautela para tratar de mantener el respeto científico que se había granjeado entre los físicos de su época. En enero de 1896 había ya escrito tres comunicaciones científicas directamente relacionadas con sus observaciones, cuyo contenido era tan minucioso que tuvo que pasar una década para poder volver a decir algo nuevo sobre los rayos X.



Fig.3. Radiografía de la mano del anatomista suizo Alfred von Kölliker.

El 23 de enero de 1896, Röntgen pronunció su única conferencia pública sobre su descubrimiento en la *Sociedad de Física y Medicina de Würzburg*, en donde, además de relevantes científicos estuvieron presentes militares de alto rango y otras muchas personalidades. La conferencia se vio interrumpida varias veces por el entusiasmo de los oyentes, y en especial tras la proyección de la radiografía de la mano del anatomista suizo Alfred von Kölliker (Fig. 3), quien propuso dar el nombre de Röntgen a esta nueva radiación, tal como se viene denominando en alemán a los rayos X (*Röntgenstrahlung*).

La respuesta científica al hallazgo de Röntgen (Fig. 4) fue colosal, y aunque Röntgen nunca ocultó que sus indagaciones se basaron en estudios de otros colegas, como los alemanes J.W. Hittorf (1824-1914) y P. Lenard (1862-1947), o el británico W. Crookes (1832-1919), no faltaron las pretensiones de algunos para atribuirse el descubrimiento. Incluso el emperador alemán Guillermo II ordenó que el físico estrella se desplazara a Berlín el 13 de enero de 1896 para conocer en primera persona la naturaleza de esos rayos misteriosos que todo lo traspasaban, y es que la sensación de poder mirar en el interior del propio cuerpo hizo que los rayos X también se convirtieran en el descubrimiento físico más popular de su época.



Fig.4. Fotografía de Wilhelm Conrad Röntgen.

Inexplicablemente, Röntgen renunció a patentar el resultado de su experimento, con lo que, en principio, en aquel tiempo cualquiera podía construir un equipo productor de rayos X, y como los peligros de esta nueva radiación aún eran completamente desconocidos, su uso se hizo totalmente popular e incontrolado, llegando incluso a usarse con fines lúdicos. Sin embargo, curiosamente, la introducción de su uso para aplicaciones médicas fue mucho más lenta de lo que uno puede imaginar, hecho que en algún momento hizo flaquear el posible futuro del descubrimiento. Debieron pasar aún bastantes años hasta que su "increíble luz" fuera reconocida como de verdadero interés médico, e incluso le reportara el máximo galardón del *primer Premio Nobel de Física de 1901* [2].

A pesar de todas las exageraciones vertidas alrededor de su persona y de la cambiante popularidad del descubrimiento en sí, Röntgen se mantuvo como un personaje modesto y solitario. Le gustaba pasar los veranos en su casa en Weilheim, en los Alpes bávaros. Amaba la naturaleza y fue un ávido montañista y cazador. En 1900 aceptó una cátedra en Munich, en donde pasó el resto de su vida. Murió el 10 de febrero de 1923, cuatro años después de su esposa.

Los primeros ecos del descubrimiento.

Tal como se ha mencionado más arriba, en un principio el futuro de aquellos misteriosos rayos no parecía excesivamente prometedor. Tuvo que pasar algo más de un largo decenio desde que Röntgen descubriera su capacidad para visualizar estructuras óseas, para que, ya comenzado el siglo XX, se iniciara el desarrollo de la tecnología necesaria para su uso en Medicina.

En efecto, el “ruido científico” alrededor del hallazgo de Röntgen fue espectacular. El primer informe sobre el descubrimiento en una revista científica británica fue una breve nota en *The Electrician* publicada el 10 de enero de 1896. El autor de dicho artículo pensaba con optimismo que todo el fenómeno parecía admitir una explicación fácil, y terminaba con un injusto comentario: *hay pocas personas a las que les guste sentarse durante dos horas para hacerse una foto que solo muestre sus huesos y los anillos de los dedos*. Dicha observación, característica de otros muchos artículos aparecidos durante los primeros tiempos del descubrimiento, se basaba en la creencia de que la aplicación práctica del hallazgo era principalmente en fotografía. Las primeras reacciones de la prensa médica británica publicadas en la prestigiosa revista médica *The Lancet*, a finales de 1896, comenzaron algo molestas y en tono grotesco, aunque semanas más tarde fueron algo más cautas, dando paso a la duda sobre la existencia de un descubrimiento posiblemente importante para la Medicina. El índice del volumen 53 de la revista *Nature*, publicado en 1896, enumera no menos de 163 artículos, cartas y notas procedentes de muchos países con comentarios sobre los rayos X. Según Glasser [3] en ese mismo año llegaron a publicarse en el mundo no menos de 1.044 libros y folletos sobre el descubrimiento de Röntgen. La primera monografía sobre el tema fue publicada en Gran Bretaña en 1896 con el título *Practical Radiography* [4] y curiosamente, quizás como consecuencia de la falta de familiaridad con la anatomía humana a la luz de los rayos X, su portada mostraba un corazón humano, pero boca abajo.

La ciencia se lo toma en serio.

Röntgen comenzó a salir definitivamente de su incómoda situación gracias a su decisivo envío a la revista británica *British Medical Journal* de una radiografía con un brazo fracturado, lo cual daba idea de la capacidad diagnóstica de su descubrimiento. Desde ese momento, las reacciones en el campo médico cambiaron radicalmente, tal como describe el excelente artículo de Posner [5].



Fig.5. Una de las unidades móviles de Curie utilizada por el ejército francés. Bibliothèque Nationale de France, Département Estampes et Photographie. <https://bit.ly/2N0swCS>.

Al comienzo de la primera guerra mundial los equipos de rayos X todavía se encontraban exclusivamente en los hospitales de las grandes ciudades, lejos de los campos de batalla en donde se tenía que atender a los heridos. La solución al problema vino de la mano de la intrépida científica Marie Curie (1867-1934), quien desarrolló el primer "coche radiológico", es decir, un vehículo que contenía un equipo productor de rayos X y un aparato fotográfico con cuarto oscuro para revelar las placas radiológicas, y que podía desplazarse hasta el campo de batalla para que los cirujanos del ejército pudieran utilizar los

rayos X y guiar así su intervenciones quirúrgicas (Fig. 5). Uno de los principales obstáculos del proyecto fue la necesidad de disponer de la suficiente energía eléctrica para producir los rayos X, pero Curie resolvió el problema incorporando una dinamo, un tipo de generador eléctrico en el diseño del automóvil, de tal modo que finalmente la electricidad requerida era producida por el motor de combustión del automóvil.

El fundamento del equipamiento para producir rayos X es relativamente sencillo (Fig. 6) se requiere un generador de alta tensión (hasta unos 50.000 voltios), que se suministra al llamado tubo de rayos X, que es en donde se produce la radiación [6]. La alta tensión se suministra como diferencia de potencial (alto voltaje) entre un filamento incandescente (cátodo), por el que se hace pasar una intensidad de corriente de bajo voltaje (aproximadamente 5 A y 12 V) y un metal puro (ánodo) como wolframio, cobre, molibdeno, etc., estableciéndose entre ambos una corriente de unos 30 mA de electrones libres. Desde el filamento incandescente (cargado negativamente) saltan electrones hacia el ánodo (cargado positivamente) provocando, en los átomos de este último, una reorganización electrónica en sus niveles de energía. Los electrones previamente excitados del material anódico que retornan a su estado previo (energéticamente más bajo) emiten ese exceso de energía en forma de rayos X, característicos de ese material (Fig. 7).

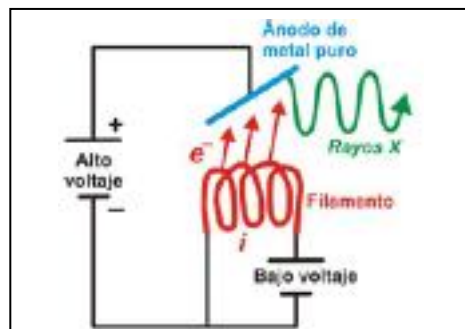


Fig. 6. Esquema de la producción de rayos X.

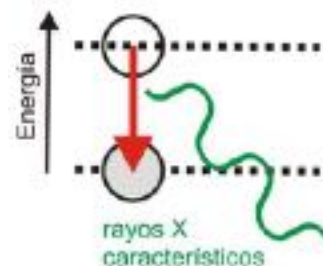


Fig.7. Los electrones previamente excitados retornan a su estado original, devolviendo esa energía en forma de rayos X.

Actualmente es de todos conocido que la radiología se ha convertido en una importantísima especialidad como ayuda para el diagnóstico médico. Es la



Fig.8. Imagen radiológica de una fractura de fémur.

herramienta fundamental para la detección de enfermedades del esqueleto, pero también se utiliza para diagnosticar enfermedades de los tejidos blandos, como la neumonía, cáncer de pulmón, edema pulmonar, abscesos, etc. Las imágenes de rayos X muestran el interior del cuerpo en diferentes tonos de grises, lo cual es debido a que los diferentes tejidos absorben la radiación de forma desigual. El calcio de los huesos absorbe la mayor parte de los rayos X, por lo que la placa fotográfica apenas se impresiona y los huesos se ven casi blancos (Fig. 8). La grasa y los restantes tejidos blandos, constituidos por elementos químicos de menor número atómico que el calcio (carbono, nitrógeno, hidrógeno), absorben menos, y se ven de color gris, y como el aire apenas absorbe la radiación, los pulmones se ven negros. En la actualidad, incluso en algunos casos, cuando el examen radiológico directo tiene limitaciones, los rayos X también dan solución

a la observación, mediante la denominada tomografía axial computarizada (TAC), sin olvidar que los rayos X también se usan para la observación de eventos en tiempo real, tales como la angiografía o, en general, estudios de contraste, administrando sustancias con cierto grado de opacidad para hacer visibles las venas, arterias o vasos linfáticos. Pero, además de su imprescindible uso como herramienta de diagnóstico, los rayos X juegan también un papel muy importante como instrumento terapéutico (radioterapia) para destruir tumores y células cancerosas al dañar su ADN, y por ello la dosis de radiación utilizada para tratar el cáncer es mucho más alta que la dosis de radiación utilizada para las imágenes de diagnóstico. La radiación terapéutica normalmente se genera en un equipo externo al cuerpo o puede provenir de un material radioactivo que se coloca en el cuerpo, dentro o cerca de las células tumorales, o inyectado directamente en el torrente sanguíneo.

Pero no sólo la Medicina ha sabido beneficiarse de las aplicaciones de los rayos X. En el mundo industrial se utilizan toda una serie de ensayos considerados como herramienta fundamental para el control de calidad de materiales (Fig. 9). Estas pruebas se basan en el distinto poder de penetración de los rayos X en los materiales, dependiendo de su naturaleza, grosor y o densidad. La finalidad es descubrir, localizar y evaluar los defectos, peculiaridades o discontinuidades, superficiales o interiores, de cualquier tipo de material, en

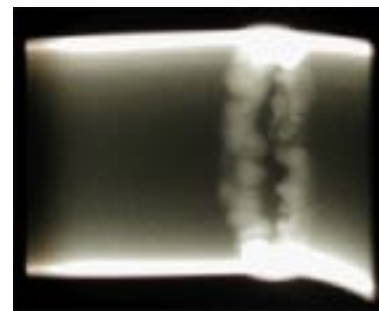


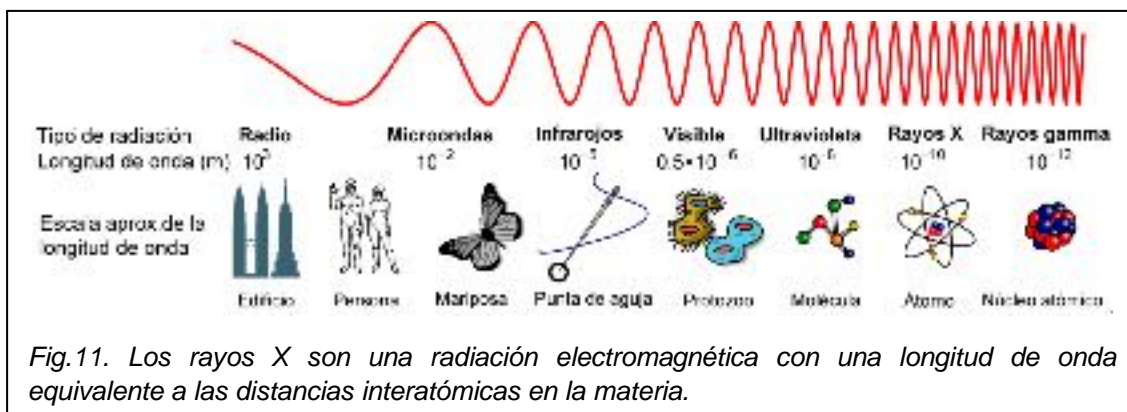
Fig.9. Imagen radiográfica de una soldadura defectuosa.



Fig.10. Una radiografía revela una pintura, previamente desconocida, de Amedeo Modigliani que surgió detrás de una de sus obras maestras. Foto de Mark Heathcote y Abbie Soanes / Galería Tate.

soldaduras, estructuras de puentes o edificios, en las estructuras de los medios de transporte, como barcos o aviones, piezas o partes de equipos, verificación de montajes, desarrollo y/o control de procesos y en cualquier otro tipo de campo industrial en donde la radiación X no afecte las propiedades o funcionalidad del material en estudio. Dentro de este tipo de ensayos podemos incluir las investigaciones que ya desde hace años se llevan a cabo en el campo del arte, como por ejemplo para la correcta reparación de pinturas famosas o en la exploración de sus orígenes (Fig. 10).

Históricamente hablando, y al margen del gigantesco desarrollo que el uso de los rayos X ha supuesto para la medicina y la industria, el mayor salto cualitativo que científicamente aportó el descubrimiento de Röntgen ocurrió el 8 de junio de 1912, cuando la Academia de Ciencias de Baviera publicó un breve comunicado firmado por Max Laue (1879-1960), entonces todavía sin el "von", y sus dos ayudantes Walter Friedrich y Paul Knipping. El comunicado anunciaba, con cierta premura, que los rayos X generaban interferencias al pasar a través de los cristales, es decir, al interaccionar con la materia ordenada, que es lo que se suponía que eran los cristales [7]. Y es que esta simple observación respondía inequívocamente a dos importantes preguntas del momento. Por un lado, el hecho de que los misteriosos rayos dieran lugar a patrones de interferencia demostraba que su naturaleza era la misma que la de una radiación electromagnética (una onda, Fig. 11), y aún más, que su longitud de onda (la distancia que recorre la onda en el intervalo de tiempo transcurrido entre dos máximos consecutivos de la onda) sería del orden de las separaciones regulares que debían de existir en el interior de los cristales, lo que, por otro lado, constataba la naturaleza ordenada de estos últimos. Una magnífica descripción histórica de estos hechos y de los correspondientes experimentos llevados a cabo bajo la dirección de Max von Laue, puede encontrarse en el artículo de Michael Eckert [8].



Aunque nadie podría restar ningún mérito al hallazgo de Laue, es justo citar que la decisión de llevar a cabo el experimento con los cristales fue consecuencia de conversaciones previas con su colega de la Universidad de Munich Paul Peter Ewald (1888-1985), quien entonces estaba acabando su Tesis Doctoral bajo la dirección de Arnold Sommerfeld (1868-1951). Ewald interesó a Laue por sus propios experimentos sobre interferencias entre radiaciones de gran longitud de onda (prácticamente luz visible) sobre un "modelo cristalino" (ordenado) basado en resonadores, y todo ello cuando

todavía estaba en discusión la naturaleza corpuscular u ondulatoria de las radiaciones. Esta idea de Ewald es la que, finalmente, llevó a Laue a imaginar qué pasaría si en lugar de tales grandes longitudes de onda se usaran otras de mucha menor longitud (mayor frecuencia), y directamente sobre los cristales quienes, en teoría, deberían comportarse como redes de interferencia muy pequeñas. Es también de justicia mencionar que tras el éxito del experimento resultó sorprendente la rapidez con la que Ewald desarrolló y publicó la interpretación matemática de las observaciones de Max von Laue [9].

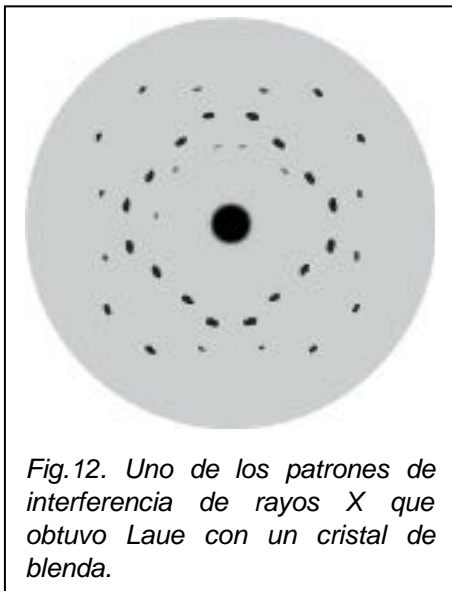


Fig.12. Uno de los patrones de interferencia de rayos X que obtuvo Laue con un cristal de blenda.

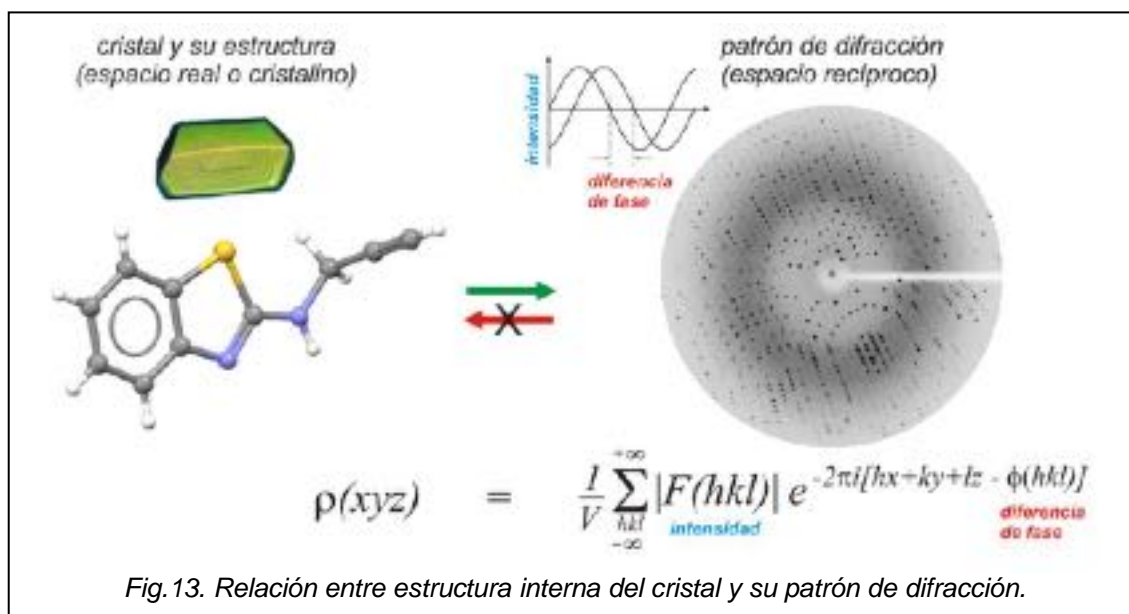
El hallazgo de Max von Laue, obteniendo los primeros diagramas de difracción con cristales (Fig.12), tuvo un eco estremecedor en el mundo científico. En Inglaterra, William Henry Bragg (1862-1942) y William Lawrence Bragg (1890-1971), padre e hijo, confirmaron el descubrimiento de Laue, y en otoño de 1912, el mismo año en que Max von Laue hizo público su experimento, el joven Lawrence Bragg comenzó a examinar el fenómeno que ocurría al interponer un cristal entre una fuente de rayos X y una placa fotográfica, presentando sus primeros

resultados en la sede de la *Sociedad Filosófica de Cambridge* en su reunión del 11 de noviembre de 1912 y publicado en enero de 1913 [10]. Muy pocos descubrimientos científicos han sido reconocidos tan rápidamente con el máximo galardón que supone el premio Nobel, ya que Max von Laue, autor del experimento de Munich, recibió el premio Nobel de 1914 [11] "por su descubrimiento de la difracción de rayos X por cristales", y los Bragg recibieron el premio de 1915 [12] "por sus servicios en el análisis de la estructura de los cristales por medio de los rayos X".

A la luz de los resultados de Max von Laue y de los Bragg (padre e hijo), los rayos X acabaron revolucionando el antiguo campo de la Cristalografía, que hasta entonces había estudiado fundamentalmente la morfología de los minerales. El fenómeno de la interacción de aquella extraña radiación con los cristales, demostró que los rayos X tenían naturaleza electromagnética, de longitud de onda del orden de 10^{-10} metros, y que la estructura interna de los cristales era discreta y periódica, en redes tridimensionales, con separaciones interatómicas de ese orden de magnitud. Estos hechos provocaron que, ya desde el pasado siglo XX, la Cristalografía se convirtiera en una de las disciplinas básicas para muchas otras ramas de la Ciencia, y en especial de la Física y Química de la materia condensada, de la Biología y de la Biomedicina.

En efecto, tal como ya demostraron los Bragg con cristales de sustancias sencillas como el cloruro sódico (sal común), el patrón de difracción originado por los rayos X al pasar por el interior de los cristales genera como la "huella digital" de su estructura interna, es decir, de la posición relativa de todos los átomos que la componen. Ello permite contemplar el experimento de la difracción como un par de escenarios conectados. El primer escenario estaría

representado por la estructura interna (átomos, moléculas, iones), y el segundo lo estaría por el propio patrón de difracción. La conexión entre ambos escenarios sería “una puerta” virtual que estaría simbolizada por el signo de igualdad (=) que, en su correspondiente formulación matemática, conecta ambos espacios (Fig. 13). El primer escenario, conocido con el nombre de “espacio real o cristalino”, está representado en dicha figura por el miembro de la izquierda de la mencionada expresión matemática; se denomina función de densidad electrónica, $\rho(xyz)$, y no contiene más complejidad que la de las posiciones de los átomos en el interior de la celdilla elemental, un paralelepípedo virtual que por apilamiento en las tres direcciones del espacio genera el volumen cristalino, tal como lo hacen los ladrillos de un muro. El segundo escenario, conocido como “espacio recíproco” (introducido por Paul Peter Ewald inmediatamente después de los experimentos de Max von Laue) está constituido por las ondas de rayos X que han sido difractadas por el cristal. Sin embargo, este escenario guarda un “cierto secreto”, pues de estas ondas difractadas solo conocemos su efecto final sobre una película fotográfica, manchas de diferente intensidad. Y es que las ondas difractadas no se caracterizan exclusivamente por sus intensidades, sino también por sus “diferencias de estado”, lo que se conoce con el nombre de “diferencias de fase” (Fig.13). Las intensidades que aparecen en el patrón de difracción son magnitudes perfectamente medibles, pero su “secreto” (las diferencias de fase entre unas ondas y otras), se nos “escapa” de la medida experimental, lo que implica una gran dificultad para movernos en el sentido de derecha a izquierda del esquema. Esta dificultad implicó que la determinación de la estructura interna de cristales más complicados que los que manejaron los Bragg no fuera conseguida en la época de los Bragg. Esta meta ha sido alcanzada gracias al esfuerzo de grandes científicos durante un centenar de años, en una época plagada de laureados Nobel, pero esa es otra historia [13, 14].



En la actualidad, gracias al conocimiento estructural que nos proporcionan los rayos X en su interacción con los cristales somos capaces de producir materiales con propiedades prediseñadas, desde catalizadores para

una reacción química de interés industrial, hasta pasta de dientes, placas de vitrocerámica, materiales de gran dureza para uso quirúrgico, o determinados componentes de los aviones, por poner algunos ejemplos. Gracias a la interacción de los rayos X con los cristales pudimos averiguar los secretos estructurales del ADN, el llamado código genético. Podemos aumentar la resistencia de las plantas frente al deterioro medioambiental. Somos capaces de comprender, modificar o inhibir, enzimas implicados en procesos fundamentales de la vida e importantes para mecanismos de señalización que ocurren en el interior de nuestras células, como el cáncer. Gracias al conocimiento de la estructura del ribosoma, la mayor fábrica

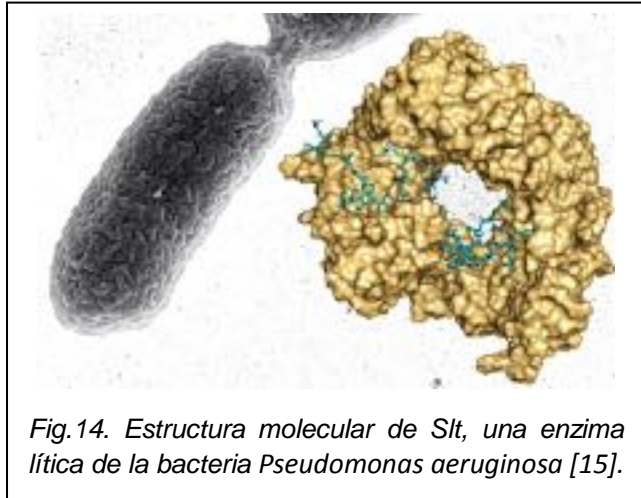


Fig.14. Estructura molecular de Slt, una enzima lítica de la bacteria Pseudomonas aeruginosa [15].

de proteínas de nuestras células, podemos entender el funcionamiento de los antibióticos y modificar su estructura para mejorar su eficacia. De la estructura de las enzimas bacterianas (Fig. 14) o producidas por ciertos virus, hemos aprendido cómo combatir los microorganismos con alta resistencia a antibióticos, y ya somos capaces de desentrañar las sutiles maquinarias de defensa que han desarrollado estos gérmenes, con lo que ya no es un sueño pensar que podemos combatirlos con herramientas alternativas a los antibióticos.

Referencias:

- [1] Röntgen, W.K. 1896. Eine neue Art von Strahlen, Verlag und Druck der Stahel'schen K.B. Hof- und Universitäts- Buch- und Kunsthandlung, Würzburg. <http://bit.ly/2wFdNCp>
- [2] The Nobel Prize in Physics 1901, Wilhelm Conrad Röntgen. Fundación Nobel. <https://bit.ly/2N0cPvB>
- [3] Glasser, O. 1933. Wilhelm Conrad Röntgen and the Early History of the Roentgen Rays. London, John Bale, Sons and Danielson. <https://bit.ly/2NwjBbK>
- [4] Ward, H.S. 1896. Practical Radiography. E.A. Robins y A.E. Livermore. London, Dawbarn y Ward.
- [5] Posner, E. 1970. Reception of Rontgen's Discovery in Britain and U.S.A., British Medical Journal, 4, 357-360, <https://bit.ly/2oEOlcQ>
- [6] The Cathode X-ray Tube site, Electronic Glassware, History and Physics Instruments, <https://bit.ly/2NqUAiw>
- [7] Friedrich, W., Knipping, P. y Laue, M. 1912. Interferenz-Erscheinungen bei Röntgenstrahlen, Sitzungsberichte der Königlich Bayerischen Akademie der Wissenschaften, Mathematische-Physische Klasse, 42, 303-322, <https://bit.ly/2Qehg3P>
- [8] Eckert, M. 2012. Max von Laue and the discovery of X-ray diffraction in 1912, Ann. Phys. (Berlin) 524, No. 5, A83–A85, [doi:10.1002/andp.201200724](https://doi.org/10.1002/andp.201200724)
- [9] Ewald, P.P. 1913. Zur Theorie der Interferenzen der Röntgenstrahlen in Kristallen. Physikalische Zeitschrift, 11, 465-472, <https://bit.ly/2MefY5R>
- [10] Bragg, W.L. 1913. The diffraction of short electromagnetic waves by a crystal. Proceedings of the Cambridge Philosophical Society, XVII (1), 43-57, <https://bit.ly/2NOUZIX>
- [11] The Nobel Prize in Physics 1914, Max von Laue. Fundación Nobel. <https://www.nobelprize.org/prizes/physics/1914/summary/>
- [12] The Nobel Prize in Physics 1915, Sir William Henry Bragg and William Lawrence Bragg. Fundación Nobel. <https://www.nobelprize.org/prizes/physics/1915/summary/>
- [13] Martínez-Ripoll, M., Hermoso, J.A., Albert, A. (coord.). 2014. A través del cristal. Cómo la cristalografía ha cambiado la visión del mundo, CSIC-Catarata, 196 págs., ISBN: 978-84-00-09800-1.
- [14] Cano, F.H. y Martínez-Ripoll, M. Cristalografía-Crystallography. Web para la enseñanza de la Cristalografía, <http://www.xtal.iqfr.csic.es/Cristalografia/>
- [15] Lee, M., Batuecas, M.T., Tomoshige, S., Domínguez-Gil, T., Mahasenan, K.V., Dik, D.A., Hesek, D., Millán, C., Usón, I., Lastochkin, E., Hermoso, J.A. and Mobashery, S. 2018. Exolytic and endolytic turnover of peptidoglycan by lytic transglycosylase Slt of *Pseudomonas aeruginosa*. PNAS, 115, 4393-4398, [doi:10.1073/pnas.1801298115](https://doi.org/10.1073/pnas.1801298115).

Bibliografía:

- (1) Bordley, J., Harvey, A. Macgehee. 1976. Two Centuries of American Medicine 1776-1976. Philadelphia, W.B. Saunders, <https://bit.ly/2MSjXdh>.
- (2) DESY 2000 - Licht der Zukunft. Wilhelm Conrad Röntgen und die Entdeckung der X-Strahlen. <https://bit.ly/2r9VEti>.
- (3) Deutsches Röntgen-Museum, <https://bit.ly/2Q9XuGT>.
- (4) Fresquet, J.L. Wilhelm Conrad Röntgen (1845-1923), en Historia de la Medicina, <https://bit.ly/2oJkK1F>.
- (5) Garrison, F.H. 1921. An introduction to the history of medicine, with medical chronology, suggestions for study and bibliographic data. W.B. Saunders Company, Philadelphia and London, <https://bit.ly/2aZlqam>.
- (6) Röntgen, W. C.. 1894. Zur Geschichte der Physik an der Universität Würzburg, Festrede zur Feier des dreihundert und zwölften Stiftungstages der Julius-Maximilians-Universität gehalten am 2ten Januar. Würzburg, Druck der KGL Universitätsdruckerei von H. Stürtz, <http://bit.ly/2LloB3c>.

Martín Martínez-Ripoll

Doctor en Ciencias Químicas.

Profesor de Investigación *Ad Honorem*, Instituto de Química-Física "Rocasolano", CSIC.

Capítulo 5

Darwin y la evolución: un gran paso para una visión objetiva de la especie humana en la naturaleza.

El genuino Darwin.

¿Qué puede saber de Darwin un lector medio, con algún interés por la biología? Probablemente tendrá unas nociones sobre el estado de la biología antes de Darwin, y de la revolución que supuso la irrupción del darwinismo, y que, hasta la llegada de Darwin dominaba el creacionismo, con una concepción de la naturaleza, creada por Dios, fija e inmutable.

La historia científica de la evolución comienza con Jean-Baptiste de Lamarck. Pero, generalmente, se insiste más en asociar a Lamarck como padre de la idea de la herencia de los caracteres adquiridos que con las auténticas señas de identidad lamarckianas: los seres vivos más primitivos surgen mediante generación espontánea y evolucionan, mediante una necesidad o impulso interno de cambio, hacia una mayor perfección. Como resultado imperfecto, de esta tendencia progresiva, se producen desviaciones adaptativas laterales frente a los cambios del entorno. En este sentido, también es muy probable que algunos lectores no sepan que Darwin (1) propuso una teoría de la herencia de carácter lamarckiano, la “*pangénesis*”, basada en la “herencia del uso y del desuso”: los hábitos adquiridos por un individuo modificarían sus órganos corporales, y éstos producirían unas entidades microscópicas, denominadas “gémulas”, que se acumularían en las gónadas, transfiriendo así las modificaciones de los órganos de los progenitores a los órganos de la descendencia. Darwin -y su teoría de la pangénesis, desacreditada en varias ocasiones- encontraría actualmente consuelo en las

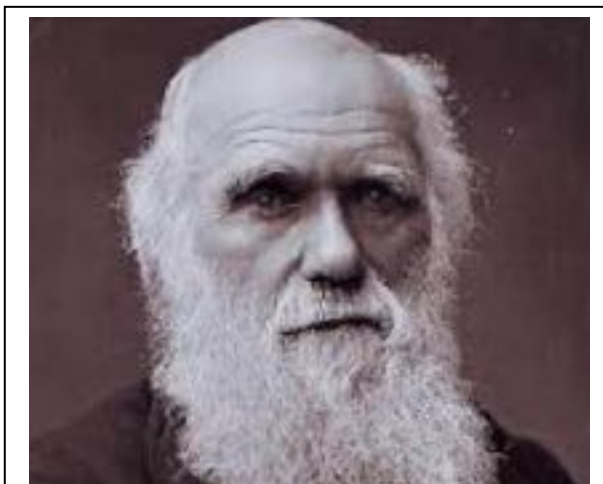


Fig. 1. Charles Darwin (1809-1882).

crecientes investigaciones sobre los exosomas: vesículas extracelulares diminutas que intervienen en la comunicación entre todos los tipos celulares, incluidos los gametos. Están cargadas de lípidos y un amplio surtido de proteínas y ácidos nucleicos, que varía en función del tipo celular y de su estado fisiológico: proteínas de adhesión celular, de fusión, transportadores de membrana, citoesqueléticas, de señalización intracelular, relacionadas con la síntesis de proteínas, de respuesta a estrés,

enzimas variadas; y, también, varios tipos de ARNs, así como múltiples fragmentos de ADN, que portarían secuencias de todos los cromosomas. García Rodríguez (2).

A diferencia de Lamarck, en la teoría darwiniana de la selección natural, la evolución se produce, sin propósito previo ni sentido alguno, mediante la generación previa de variación individual, que conlleva un aumento de las posibilidades de sobrevivir y de reproducirse –selección natural y selección sexual- de los individuos más adaptados a los cambios del medio ambiente. Darwin sabía, por la práctica de los criadores de razas domésticas, que las especies albergan una fuente inagotable de variabilidad; y que la selección natural era independiente de los mecanismos generadores de la misma.

Así, en el capítulo IV de “El origen de las especies por selección natural” Darwin (3) nos dice: “...la variabilidad que encontramos casi universalmente en nuestras producciones domésticas no está producida directamente por el hombre... Tengamos también presente cuán infinitamente complejas y rigurosamente adaptadas son las relaciones de todos los seres orgánicos entre sí y con condiciones físicas de vida. Si esto ocurre, ¿podemos dudar –recordando que nacen muchos más individuos de los que acaso pueden sobrevivir- que los individuos que tienen ventaja, por ligera que sea, sobre otros tendrían más probabilidades de sobrevivir y procrear su especie? A esta conservación de las diferencias y variaciones individualmente favorables y la destrucción de las que son perjudiciales, la he llamado yo selección natural o supervivencia de los más adecuados. Varios autores han interpretado mal o puesto reparos a la expresión selección natural. Algunos hasta han imaginado que la selección natural produce la variabilidad, siendo así que implica solamente la conservación de las variedades que aparecen y son beneficiosas al ser en sus condiciones de vida”.

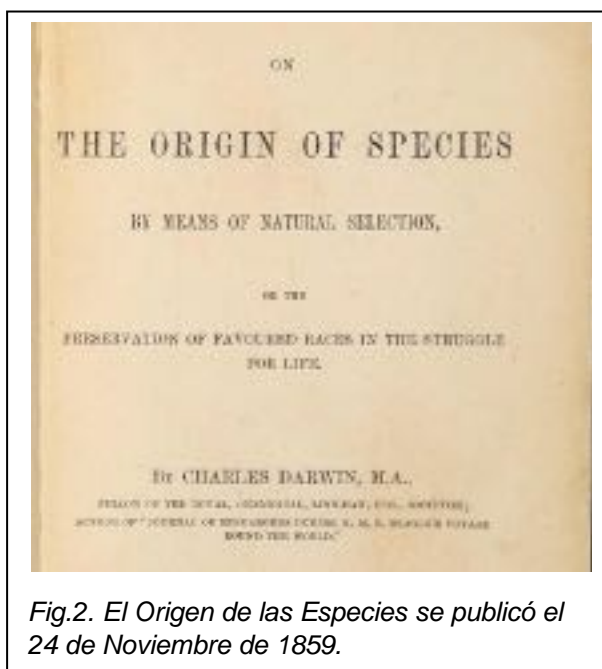


Fig.2. El Origen de las Especies se publicó el 24 de Noviembre de 1859.

Además, en el capítulo I, plantea: “Por lo general los criadores hablan de la organización de un animal como algo plástico, que se puede modelar a voluntad... Si la selección consistiera meramente en separar una variedad muy típica, y hacer cría de ella, el principio sería tan evidente como apenas digno de mención; pero su importancia reside en el gran efecto producido por la acumulación en una dirección, durante generaciones sucesivas, de diferencias absolutamente inapreciables para el ojo no experto.”

Aquí Darwin resalta la importancia de la selección en el fabuloso despliegue de formas, pero lo más importante es la llamada de atención que

nos hace de que es preciso una constelación de variaciones (“*absolutamente inapreciables para el ojo no experto*”), no una o unas pocas muy evidentes.

Así pues, a diferencia del planteamiento genético actual -presente en la teoría sintética de la evolución, pendiente de las frecuencias relativas de variantes de genes aislados que mutan- Darwin se centra directamente en el fenotipo, considerado como un todo, y en el papel del ambiente selector de ese todo. Darwin subraya el carácter conservador y acumulador –y no generador directo de variaciones- de la selección natural.

La revolución científica darwiniana.

Aún después de la exitosa publicación de *El Origen de las especies por medio de la selección natural*, Darwin se sentía incomprendido en la esencia misma de su construcción teórica. Así, en su autobiografía (1887) declara:

“Se ha dicho a veces que el éxito del Origen demostraba que ‘el tema flotaba en el ambiente’, o que ‘la mente humana estaba preparada para él’. No creo que sea estrictamente cierto, pues, de vez en cuando, sondeé a no pocos naturalistas y jamás me topé con ninguno que dudara, al parecer, sobre la permanencia de las especies. Ni siquiera Lyell o Hooker parecieron estar nunca de acuerdo conmigo, a pesar de que solían escucharme con interés”. Darwin (4).

¿De dónde viene el rechazo e incompreensión a la teoría darwiniana que, en parte, llega a nuestros días? Muchos autores coinciden en que los problemas con la obra de Darwin vienen del concepto de selección natural.

Para algunos autores darwinistas actuales el problema de aceptación de la teoría de la selección natural, tal como la formuló Darwin, es de índole filosófica cuando no religiosa. Así, en 1977, Gould (5) se pregunta: “¿Por qué ha resultado Darwin tan difícil de asimilar?” “No se impuso [la selección natural] hasta la década de 1940, e incluso hoy en día... sigue siendo ampliamente mal interpretada, se cita con errores y se aplica mal”. Gould continúa explicando que el problema radica en el planteamiento filosófico materialista de Darwin: “Darwin temía sacar a la luz algo que percibía como mucho más herético que la propia evolución: el materialismo filosófico, el postulado de que la materia es la base de toda la existencia y de que todos los fenómenos mentales y espirituales son sus productos secundarios. No existía idea alguna que pudiera resultar más demoledora para las enraizadas tradiciones del pensamiento occidental que la afirmación de que la mente –por compleja y poderosa que fuera- era un producto del cerebro...”.

A este respecto, Eldredge (6) y Gould (5) comentan las distintas posiciones filosóficas y religiosas de algunos autores, evolucionistas o no, donde se aprecia la radical diferencia con el materialismo monista de Darwin.

En primer lugar -por proximidad y méritos propios en la formulación de una teoría de la evolución por selección natural, de forma independiente a la realizada por Darwin- conviene mencionar a Wallace; subrayando,

fundamentalmente, el carácter netamente dualista de su concepción diferencial de la mente y del cerebro humanos.

Así, en su libro *El origen del hombre*, Darwin nos dice:

“Debió realizarse un extraordinario progreso en el desarrollo del entendimiento, así que entró en uso, mitad por arte y mitad por instinto, el lenguaje, pues el hábito repetido de la palabra al obrar activamente sobre el cerebro y producir efectos hereditarios, impulsaba a la vez el perfeccionamiento del lenguaje... el volumen del cerebro humano, en relación con el cuerpo, comparado con el de los animales inferiores, puede atribuirse principalmente al uso precoz de una forma simple de lenguaje. Las facultades intelectuales del hombre más elevadas, como las de raciocinio, abstracción, propia conciencia, etc., son probablemente consecuencias del constante mejoramiento y ejercicio de las otras facultades intelectuales”.

Y, más adelante, hablando de la selección sexual, añade:

“El que admita el principio de la selección sexual, se verá conducido a la notable conclusión de que el sistema nervioso no tan sólo regula la mayor parte de las funciones existentes en el cuerpo, sino que ha influido directamente sobre el progresivo desarrollo de varias estructuras corporales y de ciertas cualidades mentales...; y estas facultades del entendimiento dependen manifiestamente del desarrollo del cerebro”. Darwin (7).

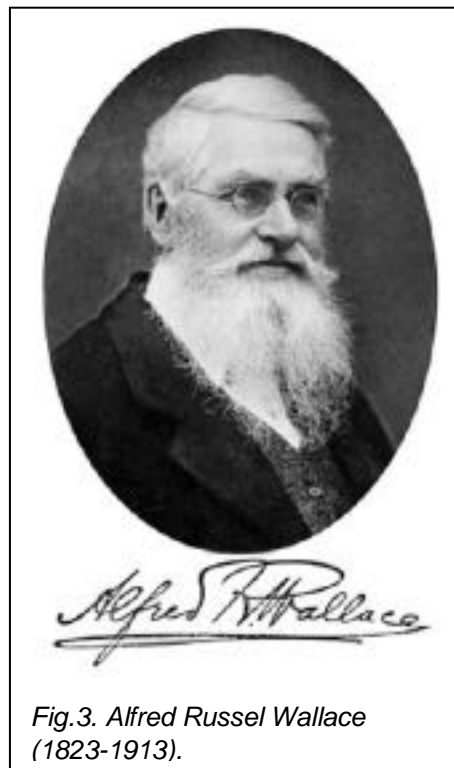


Fig.3. Alfred Russel Wallace (1823-1913).

Ambiente científico en la época de Darwin.

Continuando con el ambiente científico de la época, Eldredge (6) apunta que los intelectuales y científicos se dividían principalmente en dos grandes grupos: el de los clérigos, que dedicaban parte de su abundante tiempo libre al estudio del mundo natural, y el de los hombres con fortuna suficiente para poder dedicarse a la ciencia. Entre los que tuvieron mayor influencia en Darwin, encontramos a Adam Sedgwick y a John Stevens Henslow, ambos clérigos y profesores de universidad; y, entre los segundos, podemos destacar a Charles Lyell, abogado prestigioso que, a pesar de su fortuna familiar y personal, le dedicaba tanto tiempo y pasión a la ciencia como para dar clases en la universidad, y ser uno de los padres de la geología moderna. Lyell, en su obra *Principios de Geología*, destaca el carácter gradual de los fenómenos geológicos actuales para entender los pasados, en oposición a las explicaciones catastróficas del relato bíblico como el diluvio universal. Las posiciones gradualistas de Darwin, de las que dudaba en ocasiones, tenían

este origen geológico y, como veremos más adelante, en biología se oponían al catastrofismo de Cuvier.

Pero entre estos dos grupos de científicos, estaban emergiendo científicos de nuevo cuño: los científicos profesionales, que como profesores universitarios percibían un sueldo por su trabajo.

Darwin estableció contacto con muchos de estos científicos profesionales; y, entre los primeros estaba Robert Grant, profesor de la universidad de Edimburgo, que inició a Darwin en la metodología rigurosa de la recogida de muestras de invertebrados. También el botánico Joseph Hooker, con el que Darwin mantuvo una constante relación de amistad y de respeto mutuo, aunque no compartieran muchas de sus ideas sobre el mundo natural.

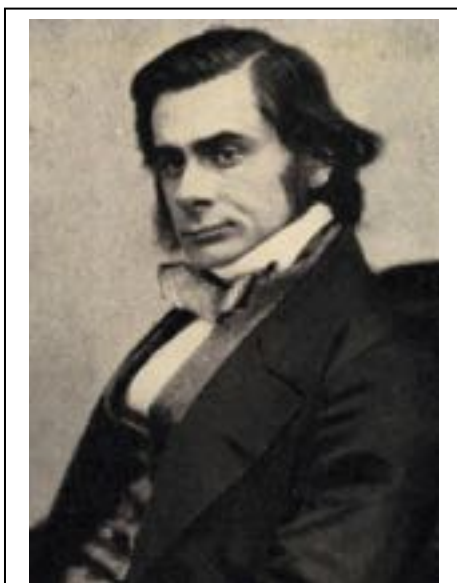


Fig.4. Thomas Henry Huxley (1825-1895).

Pero, sin duda, el científico profesional más importante para Darwin fue Thomas Henry Huxley, prestigioso profesor de anatomía comparada y gran protector de Charles, que defendió con gran convicción y fiereza *El origen de las especies*, hasta el punto de recibir el apodo de “Bulldog de Darwin”. Con la publicación del *Origen*, en 1859, Huxley encontró una nueva concepción evolucionista de la naturaleza con la que enfrentarse a su colega anatomista Richard Owen (director de la colección de ciencias naturales del Museo Británico), y, a su través, a las ideas religiosas sobre la naturaleza. Como la mayoría de los anatomistas de la época, Owen era esencialista, se oponía a la evolución biológica, y creía en la existencia de “arquetipos” anatómicos básicos creados por Dios.

Por su parte, Charles Darwin gozaba de una gran independencia económica, religiosa y política –no necesitaba trabajar para vivir, ni como clérigo ni como profesor universitario- por lo que, en principio, podría disponer de una total libertad de pensamiento; pero, como veremos, estas circunstancias le llevaron a padecer una enorme soledad: la soledad de un científico aficionado, firme defensor de sus ideas, de carácter conciliador en el trato personal, pero no acomodaticio ni condescendiente en el compromiso con su obra científica. Quizá no fuera totalmente consciente del alcance de su decisión de ser un científico independiente, al modo de Lyell, aunque Darwin no llegó a ser profesor de universidad. Recordemos que Darwin comenzó su carrera científica como geólogo, precisamente siguiendo los pasos de Lyell; pero es su paso a la biología -y, sobre todo, el descubrimiento de su teoría (“mi teoría”, como él la llamaba) la selección natural- el que marca su posición diferencial con el resto del mundo científico.

Circunstancias vitales que forjaron la obra de Charles Darwin.

Pero, ¿cómo era Darwin? ¿Cómo era ese genio que dio un giro copernicano a la forma de ver la naturaleza, incluida la naturaleza humana como un producto más en ella? ¿De dónde surgió tanto talento?

El mismo Darwin, en su autobiografía (4), agrupa sus recuerdos alrededor de tres etapas, destacando la importancia central del viaje del Beagle en el desarrollo de su carrera científica.

La etapa de formación inicial –previa al viaje del Beagle- es donde Darwin analiza las características heredadas de sus padres: sus capacidades mentales congénitas y su temperamento, junto con los recuerdos, principalmente familiares y académicos, de las circunstancias que le llevaron a modelar inicialmente su mente y su carácter.

Charles Darwin nació el 12 de febrero de 1809 en el seno de una familia acomodada, y recuerda a su padre como “el hombre más cariñoso que he conocido”, y como “el hombre más grande que he visto”; pero, también le impresionaban, y mucho, su inteligencia y su enorme capacidad de observación, y, quizá por todo esto, Darwin temía defraudar a su padre, que era médico, al igual que su abuelo Erasmus –a su vez, también naturalista y poeta-; y, aunque Charles parecía estar abocado a seguir la carrera de medicina, él dudaba de sus capacidades para ello. De entrada, quizá acomplejado por las brillantes cualidades paternas, cuestionaba su propia capacidad mental: “Mi padre, según le oí decir, creía que los recuerdos de las personas de mente poderosa se remontan, en general, muy atrás, hasta periodos muy tempranos de su vida. No es mi caso...”. “Antes de asistir al colegio fui educado por mi hermana Caroline... Me han contado que era mucho más lento para aprender que mi hermana menor, Catherine, y creo que fui en muchos sentidos un chico travieso”. El primer colegio de Charles fue sin internado, y de esa época él destaca su “gusto por la historia natural” y “la pasión por coleccionar...” “me sentía interesado, al parecer, ¡por la variabilidad de las plantas!”. El segundo colegio de Charles, en régimen de internado, fue el del Dr. Butler, también en Shrewsbury; donde permaneció siete años, hasta los 16, sin gran provecho: “Nada pudo haber sido peor para mi desarrollo intelectual...” “Cuando deje el colegio no era ni avanzado ni retrasado para mi edad; creo que todos mis maestros y mi padre me consideraban un muchacho corriente, más bien por debajo del nivel intelectual normal”. Pero lo que más le mortificaba era una frase que le espetó



Fig.5. El niño Charles Darwin (1816).

su padre: “*Lo único que te interesa es la caza, los perros y cazar ratas, y vas a ser una desgracia para ti y para toda tu familia*”.

Quizá para entender todo el proceso de la enorme proeza de Darwin, convenga saltar al final de la historia. Ya en 1876, seis años antes de su muerte, Darwin escribe su autobiografía (4) y, al final de la misma introduce un capítulo que lleva por título “*Valoración de mis capacidades mentales*”. A pesar de que, en ese momento, ya había publicado lo principal de su obra, y gozaba de gran prestigio y reconocimiento en el mundo científico, Darwin sigue viéndose como una persona poco brillante en sus capacidades intelectuales: “No soy consciente de que mi mente haya cambiado durante los últimos 30 años”. “Sigo teniendo tanta dificultad como siempre para expresarme con claridad y concisión..., pero que, como compensación, ha tenido la ventaja de obligarme a pensar largo y tendido cualquier frase...”. “No poseo una gran rapidez de entendimiento o de ingenio, tan notable en algunas personas inteligentes, como, por ejemplo, en Huxley”. “Mi capacidad para el pensamiento prolongado y puramente abstracto es muy limitada; además, nunca habría tenido éxito en el terreno de la metafísica o las matemáticas. Mi memoria es amplia pero imprecisa...”.

Aun admitiendo algún grado de modestia en las opiniones de Darwin, no podemos hablar de falsa modestia; su sinceridad y honradez intelectual están fuera de toda duda. En este sentido, Darwin también considera que posee algunas capacidades notables: “Como saldo a favor, pienso que soy superior al común de los mortales para percatarme de cosas que no atraen fácilmente la atención y observarlas con cuidado. Mi diligencia en observar y recabar datos ha sido casi todo lo grande que podía ser...mi amor por la naturaleza ha sido siempre constante y ardiente... Desde mi primera juventud he experimentado un deseo fortísimo de entender o explicar todo cuanto observaba –es decir, de agrupar todos los datos bajo leyes generales-. Todas estas causas unidas me han proporcionado la paciencia para reflexionar o sopesar durante varios años cualquier problema inexplicado. Hasta donde puedo juzgar, no estoy hecho para seguir ciegamente la guía de otras personas. Me he esforzado constantemente por mantener mi mente libre...”. Y, ya en la última página, concluye: “Por tanto, independientemente del nivel que haya podido alcanzar, mi éxito como hombre de ciencia ha estado determinado, hasta donde me es posible juzgar, por un conjunto complejo y variado de cualidades y condiciones mentales. Las más importantes han sido el amor a la ciencia, una paciencia sin límites al reflexionar largamente sobre cualquier asunto, la diligencia en la observación y recogida de datos, y una buena dosis de imaginación y sentido común. Es verdaderamente sorprendente que, con capacidades tan modestas como las mías, haya llegado a influir de tal manera y en una medida considerable en las convicciones de los científicos sobre algunos puntos importantes”.

¿Es realmente tan sorprendente que Darwin lograra explicar lo que John Herschel denominó “el misterio de los misterios”? En este sentido, las críticas negativas de Herschel, y la de otros científicos amigos de Darwin, a la selección natural, le produjeron un cierto desánimo. Por otra parte, el incondicional Huxley, exclamó al escuchar la formulación de la teoría de la selección natural: “¡Qué increíblemente estúpido no haber pensado en ello!”.

Pero, ¿por qué es tan peligrosa la teoría de la selección natural como para provocar oleadas de indignación y desagrado, incluso en nuestros días?

Para abordar estas preguntas vamos a retroceder a algunos aspectos de la etapa de formación académica de Charles, concretamente a su tendencia innata al coleccionismo y la clasificación.

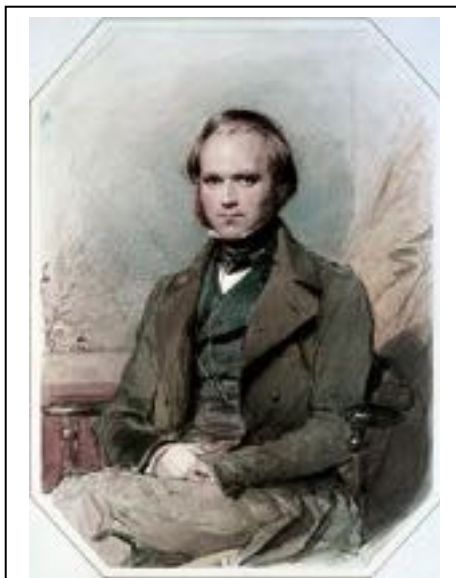


Fig.6. El joven Charles Darwin (1830).

El gran fracaso en los estudios de medicina, en la Universidad de Edimburgo, se vio compensado por algunos contactos que Darwin realizó allí, y que le permitieron perfeccionar y profundizar sus conocimientos sobre recolección, tratamiento y clasificación de especímenes biológicos. También le fueron muy útiles unas clases de pago que recibió para aprender a disecar animales. Estos nuevos conocimientos le resultaron muy valiosos para su futura dedicación a la ciencia; así que, cuando su padre decidió mandarle a Cambridge a estudiar Teología, estableció allí nuevos contactos que le permitieron profundizar en su formación de naturalista. En los cuatro años que Darwin pasó en Cambridge (1828-1831), los dos profesores más decisivos para su futura carrera científica, fueron el botánico J. S.

Henslow y el geólogo A. Sedgwick, ambos clérigos. Este profesor de botánica, con el que Darwin daba largos paseos, fue un verdadero amigo para él, y decisivo en su vida. Muy pronto le introdujo en su vida familiar, y llegó a darle alojamiento en su casa; pero su intervención más importante fue su recomendación para viajar en el *Beagle*, como naturalista no retribuido y alojado en el camarote del capitán Robert Fitz-Roy. Henslow tomó la decisión de implicar a Darwin en este viaje, porque estaba seguro de las capacidades de Darwin como naturalista, y porque también sabía que no era hombre de Iglesia; pero lo que él no podía intuir era el salto prodigioso que iba a dar la mente de este joven coleccionista apasionado por la naturaleza; Darwin tampoco. El *Beagle* comenzó su singladura el 27 de diciembre de 1831; en él iba un nuevo Charles Darwin, lleno de dudas, pero dispuesto a tensar al máximo su nueva libertad y su amor por la naturaleza.

El viaje de circunnavegación del *Beagle* duró cinco años (1831-1836) y, como el mismo Darwin reconoce sería como un nuevo nacimiento, el comienzo de su segunda vida. Dejaba atrás preocupaciones y miedos, quizá el mayor el de defraudar las expectativas de su padre. Ahora, a sus veintidós años, tenía el mundo natural por descubrir; y, en mayor o menor grado, la confianza de sus profesores y familiares. Esta confianza se asienta en las mismas cualidades positivas que Darwin enumera en su autobiografía al final de sus días. Esas cualidades que a Darwin le parecían de poco lustre para alcanzar la fama que

todos los lugares visitados fue mucho más importante, pues es en ella donde se pone en juego el razonamiento”.

Durante el viaje, Darwin estudió a Lyell, admirando la superioridad de sus argumentos sobre la del resto de los geólogos de la época, fundamentalmente su gradualismo, en oposición a los planteamientos catastrofistas de fijistas y creacionistas. Las primeras contribuciones científicas de Darwin fueron en este campo; y, así, describió y explicó la geología de Santiago, elevaciones y hundimientos que afectaban a volcanes, los orígenes y los efectos de los terremotos, también resolvió el problema de las islas de coral, entre otras aportaciones: “Fue entonces cuando caí en la cuenta por primera vez de que, quizá, podía escribir un libro sobre la geología de los diversos países visitados por mí, lo que me hizo estremecer de placer”. Así, en 1842 se publicó *La estructura y distribución de los arrecifes de coral*; en 1845 la nueva edición, que corrige la de 1839, del *Diario de investigaciones*, donde relata sus impresiones del viaje del Beagle; y en 1846 se publica *Observaciones geológicas sobre Sudamérica*.

Por su parte, la biología le puso más dificultades para elevar sus conocimientos a teoría, deslumbrado por la exuberancia de aquella naturaleza: “El esplendor de la vegetación de los trópicos...la sensación de sublimidad que me producían los grandes desiertos de la Patagonia y las montañas de la Tierra del Fuego, cubiertas de bosques...La visión de un salvaje desnudo en su tierra nativa...Muchas de mis excursiones a caballo por territorios agrestes o en barca, algunas de las cuales duraron varias semanas...”. Inicialmente, en lo relativo a los seres vivos, Darwin observó, coleccionó, describió y clasificó, mejorando notablemente estas capacidades suyas con la práctica, pero lo hizo fascinado y abrumado por la lujurante naturaleza que le rodeaba. No obstante, de forma imperceptible, Darwin iba tejiendo su singular trama de capacidades y experiencias que harían de él un gran científico. Así, en relación al viaje, le concede una gran importancia al “hábito adquirido entonces de una enérgica laboriosidad y una atención intensa en todo cuanto emprendía. Procuraba que cualquier cosa sobre la que pensaba o leía influyera directamente en lo que había visto o era probable que viese; y mantuve ese hábito intelectual durante los cinco años de viaje. Estoy seguro de que fue ese entrenamiento lo que me ha permitido hacer todo cuanto he llevado a cabo en ciencia”.

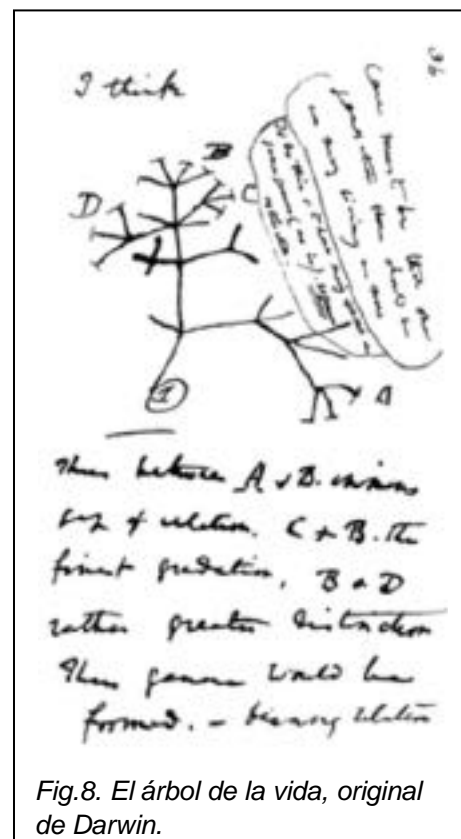


Fig.8. El árbol de la vida, original de Darwin.

La transformación que, en el viaje, estaba experimentando Darwin era notable. A diferencia de lo que decía cuando realizó, con Sedgwick, la excursión geológica al norte de Gales, ahora la ciencia ocupaba el primer lugar

en su cabeza: “mi amor por la ciencia se impuso gradualmente a cualquier otro gusto..., la caza constituía un obstáculo para mi trabajo...”.

Pero, volviendo a la biología, es en las Galápagos y, posteriormente, en Australia donde Darwin empieza a ver algo de luz entre tanta espesura: “el descubrimiento de las singulares relaciones entre los animales y plantas que poblaban las diversas islas del archipiélago de las Galápagos y las existentes entre todos ellos y los que habitan América del Sur”. Al final del viaje Darwin tiene ya una clara problemática biológica: concede una importancia a la distribución geográfica de las especies en el continente y en las islas, de manera que comienza a pensar en la transformación o modificación de unas especies en otras; y quizá barruntara también algo acerca del origen animal del hombre. Pero es en las Galápagos donde Darwin comienza a vislumbrar el *árbol de la vida*.

Darwin regresa a Inglaterra el 2 de octubre de 1836, y es un hombre completamente distinto del que partió cinco años antes.

Poco después de su regreso, comenzó a explicar los hechos observados en las Galápagos, relacionándolos con la evolución divergente: en el continente había una especie de pinzón que fue el antecesor común de las diferentes especies de pinzones que poblaban cada isla, adaptadas a nichos ecológicos distintos, en la misma o en distintas islas.

Pronto empiezan sus primeras publicaciones, aunque, por distintos motivos, se resiste a publicar su teoría principal (“mi teoría” como decía él), la selección natural: “No tardé en constatar que la selección era la clave del éxito del ser humano en la creación de razas útiles de animales y plantas. Pero durante un tiempo fue para mí un misterio cómo se podía aplicar la selección a organismos que vivían en estado natural.

En octubre de 1838, es decir, 15 meses después de haber iniciado mi indagación sistemática, leí por casualidad y para entretenerme el libro de Malthus *Sobre la población*, y como, debido a mi larga y continua observación de los hábitos de los animales y las plantas, me hallaba bien preparado para darme cuenta de la lucha universal por la existencia, me llamó la atención enseguida que, en esas circunstancias, las variaciones favorables tenderían a preservarse, y las desfavorables a ser destruidas. El resultado de ello sería la formación de nuevas especies. Ahí tenía, por fin, una teoría con la que trabajar; pero me preocupaba tanto evitar cualquier prejuicio que decidí no escribir durante un tiempo ni siquiera el menor esbozo de la misma”. Darwin (4).

¿Por qué tarda Darwin más de veinte años en publicar “su teoría” de la selección natural?

Muchos autores creen que Darwin padecía una neurosis, propiciada, en parte, por su lucha interna al tener que elegir entre ser uno de los mejores naturalistas de su tiempo, o ser honesto con sus sorprendentes descubrimientos, y asumir

la responsabilidad de iniciar una auténtica revolución en las ciencias naturales. Pero, ¿cómo guardarse ese descubrimiento? ¿Cómo disimularlo? Darwin evita entrar en liza con la religión; por prudencia pero también por respeto, fundamentalmente hacia familiares y amigos con sentimientos religiosos.

No obstante, su teoría supuso un gran paso para una visión objetiva de la especie humana en la naturaleza viva; por fin liberada de la creación divina, y dueña y responsable de su propio destino.

Charles Darwin nos demostró que la realidad se nos ofrece fácilmente cuando no queremos violentarla: basta con observar, describir, ordenar, relacionar, experimentar y, sobre todo, buscar la coherencia de la realidad con honestidad, con honradez intelectual, con planteamientos objetivos. Pero sí, por el contrario, colocamos en el centro del problema lo que no es, entonces todo se retuerce y se complica con esferas armilares y creaciones divinas.

La gran curiosidad y la honestidad intelectual de Darwin constituyeron la brújula que le permitió andar por la naturaleza sin prejuicios y sin intereses bastardos, alejado de soluciones acomodaticias y contemporizadoras. Sólo así consiguió desvelar algunos de los misterios de la vida que todos tenían ante sus ojos.

Bibliografía:

- (1) Darwin, C. (2008). *La variación de los animales y las plantas bajo domesticación*. Editorial Catarata. Madrid.
- (2) García Rodríguez, A. (2018). *¿Me conoces? Soy un exosoma*. UAM Ediciones. Madrid.
- (3) Darwin, C. (1980). *El origen de las especies*. Ed. Bruguera. Barcelona.
- (4) Darwin, C. (2009). *Autobiografía*. Editorial Laetoli. Pamplona.
- (5) Gould, S. J. (2010). *Desde Darwin. Reflexiones sobre historia natural*. Crítica. Barcelona.
- (6) Eldredge, N. (2009). *Darwin. El descubrimiento del árbol de la vida*. Katz Editores. Buenos Aires. Madrid.
- (7) Darwin, C. (2004). *El origen del hombre*. Ed. EDAF. Madrid.
- (8) <http://estructuraeinformacionbiologica.blogspot.com/>

Alfonso Ogayar Serrano

Biólogo y Profesor de Enseñanza Secundaria.

Capítulo 6

El cálculo infinitesimal: llave de la mecánica en el fecundo siglo XVII.

Explicar el cálculo infinitesimal como un gran paso de la humanidad requiere presentarlo como un devenir. Su pasado se aprecia como una acumulación de avances de diverso alcance realizados aquí y allá por la comunidad matemática (entendida como corresponde a cada época). Llegado el momento, cristalizan en el lugar más abonado como un presente creativo protagonizado por las mentes más geniales que, meditando sobre el material acumulado, ven más y mejor que quienes les precedieron y sus contemporáneos. El presente creador del cálculo infinitesimal se produjo en el último tercio del siglo XVII y los colosos protagonistas fueron el inglés Isaac Newton (1642-1727) y el alemán Gottfried W. Leibniz (1646-1716).

En este breve ensayo no será posible entrar en detalles, comparar las ideas clave de Newton (cinemáticas) con las de Leibniz (combinatorias), ni mucho

menos en dar noticia puntual de la sucesión temporal de los logros de cada uno y de la polémica sobre prioridad que los británicos provocaron.

Tras los fundadores, enseguida llegó un futuro en el que las aplicaciones del cálculo infinitesimal justifican que se le califique como un gran paso de la humanidad, que fue la llave para otros muchos progresos en diferentes ámbitos del conocimiento y la práctica.

Sin minusvalorar los avances del siglo anterior, el siglo XVII fue decisivo para la ciencia en Europa, se produjeron cambios profundos en la visión del mundo y en la capacidad humana para intervenir en la naturaleza. Avances técnicos y logros teóricos se realimentaron para llevar a cabo una revolución científica. El estudio científico del movimiento fue uno de los logros cruciales. El movimiento de los planetas, de los proyectiles, de las máquinas, alcanzó una formulación matemática gracias a una nueva herramienta: el cálculo infinitesimal. No solo hizo posible la mecánica, también propició los estudios del cambio continuo en todo tipo de problemas físicos, de la probabilidad, la economía, la biología, etc.



Fig.1. Retratos de Newton (izda) y Leibniz (dcha).

Primeras preguntas.

¿Qué es el cálculo infinitesimal? Habrá que ir contestando a esta pregunta por aproximaciones sucesivas y considerando sus partes más sencillas que permitan apreciar su esencia sin demasiados tecnicismos. Vendrá bien proceder con esa visión genética que sostiene que para entender mejor conviene no perder de vista el proceso histórico.

¿Qué es el cálculo? Un cálculo es una piedra, los pastores antiguos ponían piedras en una cesta para contar ovejas. Distribuyendo piedras en diversos recipientes se realizaban cálculos numéricos (sumas, restas, etc.), para los que el instrumento preferido fue durante muchos siglos el ábaco. Cuando llegó el álgebra, las ecuaciones se resolvían haciendo cálculos (cuentas) con los datos conocidos hasta obtener el valor de la incógnita. Estos cálculos numéricos y algebraicos se hacían con cantidades finitas y mediante un número finito de operaciones con ellas.

¿Qué significa infinitesimal? Lo infinitesimal es lo infinitamente pequeño, lo inverso de lo infinitamente grande. Sumando 1 sucesivamente con mucha paciencia, los números n se van haciendo infinitamente grandes, mientras que sus fracciones inversas $1/n$ se van haciendo infinitamente pequeñas. Algunas sumas de sumandos que se van haciendo infinitamente pequeños crecen hasta un tope, como los inversos de las potencias sucesivas de 2:

$$\frac{1}{2} + \frac{1}{4} + \frac{1}{8} + \frac{1}{16} + \dots + \frac{1}{2^n} + \dots$$

Los matemáticos no dudan en dar el valor 1 a esta suma infinita, porque si primero recorro la mitad del camino, luego avanzo la mitad de lo que queda y así sucesivamente, está claro que me acercaré a 1 todo lo que quiera y en ese infinito final llegaré a 1. Los griegos clásicos ya conocían esta suma y otras análogas con su *método de exhaustión*, que utiliza una reducción al absurdo: para demostrar que una suma infinita es igual a S se prueba que no puede ser menor que S ni mayor.

Por aquí empieza el cálculo infinitesimal, sumando infinitas cantidades que se van haciendo infinitamente pequeñas. Las partes de la magnitud

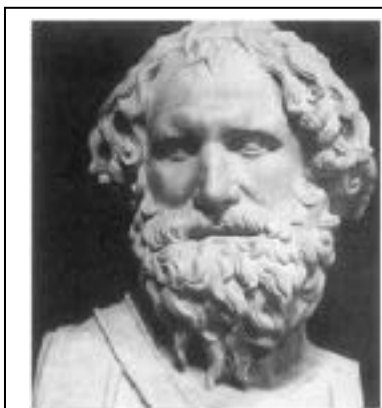


Fig.2. Arquímedes de Siracusa (s. -III).

discreta se cuentan con una unidad indivisible, pero en las magnitudes continuas hay posibilidad ilimitada de crecer y de dividirse. La esencia del cálculo infinitesimal está en el estudio del continuo, del espacio y del tiempo de la geometría como sustrato de física (con otras magnitudes continuas como el peso, la temperatura, la presión, etc.).

Arquímedes.

El gran Arquímedes usó una suma infinita para *cuadrar* un segmento de parábola, que es el

área entre la parábola y una cuerda. Cuadrar significa comparar un área con otra.

Arquímedes encontró la razón entre el área S del segmento y el área T de un triángulo asociado: cada cuerda determina una tangente paralela y un triángulo de lado la cuerda y vértice opuesto el punto de tangencia. Este fue su argumento: El triángulo deja otros dos segmentos, uno a cada lado, cada uno con su triángulo que, por propiedades de la parábola, resultan tener áreas iguales a $T_1 = \frac{1}{8}T$. El área del primer triángulo T más los dos T_1 vale $T + \frac{1}{4}T = \left(1 + \frac{1}{4}\right)T$.

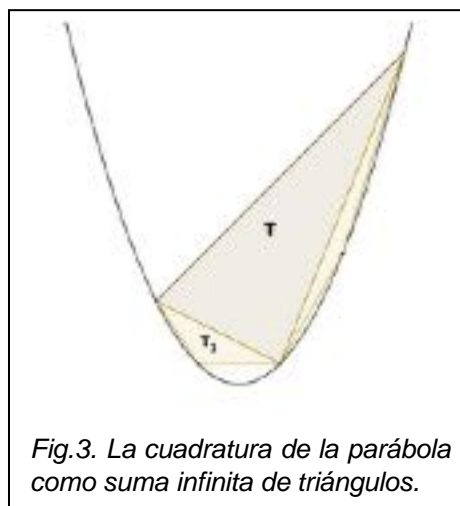


Fig.3. La cuadratura de la parábola como suma infinita de triángulos.

Para completar S quedan cuatro segmentos cada uno con su triángulo: resultan de nuevo todos iguales a $T_2 = \frac{1}{8}T_1 = \frac{1}{8^2}T$. Las áreas de los siete triángulos que van llenando S suman $\left(1 + \frac{1}{4} + \frac{1}{4^2}\right)T$. Continuando indefinidamente la suma de triángulos hasta dejar “exhausto” S , Arquímedes alcanza la suma ilimitada

$$S = \left(1 + \frac{1}{4} + \frac{1}{4^2} + \frac{1}{4^3} + \dots + \frac{1}{4^{n^n}} + \dots\right)T = \frac{4}{3}T.$$

El método que usó para llegar a $\frac{4}{3}$ tuvo una parte heurística con inspiración en la estática (ley de la palanca) y luego una prueba matemática del resultado por el método de exhaustión. Arquímedes hizo una segunda demostración por un método de aproximación por defecto y por exceso que usaron sus seguidores del siglo XVII.

Otras características de los “métodos infinitesimales” aparecen en el estudio de la espiral, curva que el sabio de Siracusa definió mediante movimientos: un segmento OA de longitud R gira alrededor de O con movimiento uniforme, de modo que A describe una circunferencia Γ de centro O y radio R . Al mismo tiempo, un punto P se mueve en el segmento OA partiendo de O con movimiento uniforme para ir hasta A mientras OA da la primera vuelta. En cada punto P de la espiral, el radio vector OP tiene una longitud r igual al arco θ girado medido en radianes (ecuación $r = \theta$ en coordenadas polares); al completar la vuelta se alcanza el valor $R = 2\pi$. Usando ideas que se consagraron como intuiciones básicas sobre lo infinitesimal, Arquímedes resolvió dos problemas con esta curva:

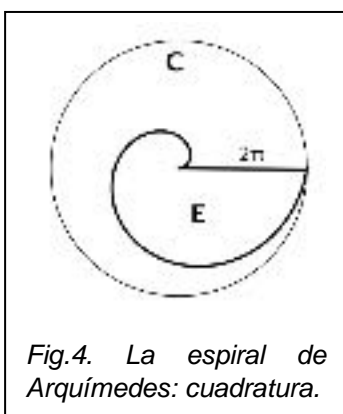


Fig.4. La espiral de Arquímedes: cuadratura.

- Cuadrar el área barrida por el segmento OP en la primera vuelta.
- Trazar la tangente en cada punto.

Comparó el área E limitada por la espiral y el segmento OA con el área C del círculo limitado por Γ . Tomó el área E como suma de sectores desde el vértice O hasta arcos “muy pequeños” de la espiral, que pudo acotar mediante sectores circulares de radios los de los extremos del arco de espiral. Dividiendo la vuelta completa en n partes iguales, el sabio de Siracusa encontró una acotación

$$\frac{s(n)}{n^3} < \frac{E}{C} < \frac{S(n)}{n^3},$$

usando una suma que sabía calcular,

$$S(n) = 1^2 + 2^2 + 3^2 + \dots + n^2 = \frac{1}{6}n(n+1)(2n+1),$$

y su versión corta $s(n) = S(n-1)$. Ambas cotas verifican

$$\frac{s(n)}{n^3} < \frac{1}{3} < \frac{S(n)}{n^3}$$

y la diferencia entre ellas es cada vez menor cuando n crece y la partición se hace más fina, así que en la posición última será

$$\frac{E}{C} = \frac{1}{3}.$$

La cuestión del trazado de la tangente es aún más interesante. Los griegos tenían una noción intuitiva de tangente: la recta que “solo toca” a la curva en un punto. Con esta idea, la tangente a la circunferencia en un punto es la perpendicular al radio, así quedó explicado en *Elementos*, la gran obra de Euclides (s. -IV/III). Más avanzado, Apolonio (s. -III) mostró cómo trazar las tangentes a las cónicas, combinando tangentes y normales con cuestiones de distancias máximas y mínimas.

Por su parte, Arquímedes halló la tangente en un punto P de su espiral dando un incremento ε “muy pequeño” al ángulo θ , lo que lleva a un punto Q de la espiral “muy próximo” a P (en la Fig. 5 está exagerada la separación). Se tiene por otra parte el arco PR correspondiente a ε pero tomado en la circunferencia de radio $OP = r$, así que $PR = r\varepsilon$. El triángulo PRQ tiene un lado recto $QR = \varepsilon$ y los otros dos curvos, siendo rectángulo en R y con ángulo α en P (los ángulos formados por las curvas son los de sus tangentes) así que la perpendicular a OP por O cortará a la tangente a la espiral en P formando ángulo α en un punto T con el segmento OT de longitud t . Arquímedes aplica la semejanza de los triángulos QRP y POT para obtener una razón de arcos igual a una razón de segmentos: $\frac{QR}{RP} = \frac{PO}{OT}$. Este

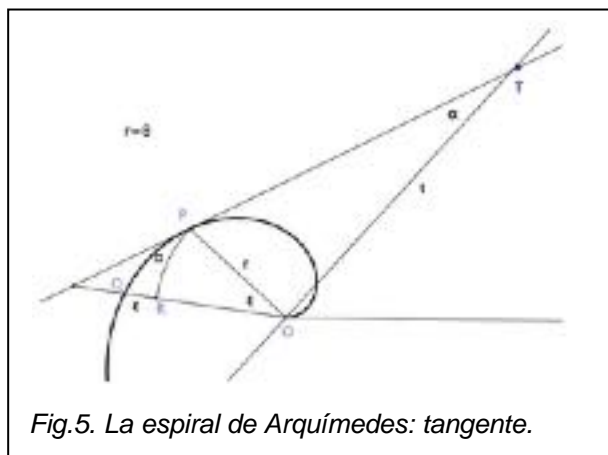


Fig.5. La espiral de Arquímedes: tangente.

argumento es muy osado, porque uno de los triángulos es curvilíneo, pero Arquímedes se siente autorizado al ser ε “muy pequeño”, de modo que el arco es “casi recto”, es “casi igual” al segmento de la tangente. Resulta pues $\varepsilon/r\varepsilon = r/t$, es decir, $t = r^2$, y la tangente PT queda determinada.

Así, con un argumento bastante atrevido, gracias a una intuición acertada sobre lo infinitesimal, logró un resultado que evaluamos como correcto con los métodos exactos actuales.

Arquímedes usó también estas habilidades para comparar volúmenes de cilindros, conos y esferas, pero la matemática griega no pudo avanzar más en esta dirección que tan brillantemente inició el siracusano. Aquella matemática fue apta para la estática pero no para el movimiento, hubo que esperar hasta que una acumulación de nuevos saberes propiciara la aparición del cálculo infinitesimal.

Tiempo de espera fructífero.

Tras el esplendor de la Grecia helenística, el periodo medieval europeo fue escaso en nuevas especulaciones, pero muy rico en experiencias prácticas. El saber griego también se transmitió a Oriente y desde allí llegaron nuevas aportaciones decisivas a través de los árabes, que descargaron su riqueza científica en el norte de África y en la Península Ibérica. Las ciudades-estado italianas usaron las rutas comerciales con Oriente para importar también conocimiento.

En los comienzos del siglo XVII, Galileo inició el estudio científico experimental del movimiento. Calcular tangentes y áreas tiene que ver con el movimiento, basta pensar que las curvas son las trayectorias por las que circulan los puntos que estudia la cinemática y las tangentes marcan en cada punto de la curva la dirección instantánea del movimiento. Torricelli, Cavalieri, Mengoli, Roberval, seguidores de Galileo, se inspiraron en los métodos de Arquímedes para avanzar en la determinación de tangentes, áreas y volúmenes, calculando sumas infinitas y aplicando composición de movimientos en las curvas engendradas por algún tipo de mecanismo.

Cavalieri imaginó que un área es engendrada por un segmento de extremos variables que se mueve paralelo a sí mismo y que todas esas líneas eran las partes indivisibles de la superficie (como el punto es la parte indivisible de una longitud), de modo que el área era la “suma de todas las líneas” y áreas “con las mismas líneas” (o proporcionales) eran iguales (o proporcionales). Lo mismo sucede con los volúmenes, son “sumas de áreas”. Con estas ideas, habilidad intuitiva y no poca destreza en los argumentos y cálculos, obtuvo una cantidad notable de relaciones entre áreas y volúmenes.

El vínculo de las áreas con el movimiento tuvo un exponente magnífico en la segunda ley de Kepler (1609): Cada planeta se mueve siguiendo una elipse en uno de cuyos focos está el Sol y lo hace con velocidad variable, de manera que permanece constante el área barrida por el segmento que lo une con el Sol en intervalos iguales de tiempo. Las leyes de Kepler surgieron de

una inducción genial a partir de las meticulosas observaciones tabuladas por Tycho Brahe, al mismo tiempo que Galileo inventaba el telescopio que mejoraría las observaciones. La segunda ley se refiere a velocidades medias, espacios recorridos en tiempos determinados, pero no a velocidades instantáneas, concepto reservado al cálculo infinitesimal que el propio Kepler vislumbró y ayudó a descubrir.

Los matemáticos posteriores objetaron la falta de homogeneidad de las sumas de Cavalieri: no se puede, decían, sumar líneas para obtener áreas, para obtener áreas hay que sumar áreas, aunque sean “muy pequeñas”, partes “infinitesimales”. En esta cuestión se aprecia, como sucedió en la Grecia clásica, que tras la matemática hay una filosofía subyacente.

Galileo dijo que el universo estaba escrito en lenguaje matemático cuando este se iba diferenciando del de los griegos, pues aparecieron dos nuevos lenguajes, uno para la aritmética y otro para el álgebra, que encontraban en la imprenta un enorme cauce de difusión nunca conocido antes:

- El sistema de numeración decimal.
- El lenguaje simbólico del álgebra.

El sistema decimal importado de Oriente permitía calcular mediante algoritmos eficaces y llevó a sustituir el ábaco por las primeras máquinas de calcular, aparecidas en el siglo XVII gracias a Pascal y a Leibniz, sin olvidar a los artesanos que las construían. Los inconmensurables que torturaron a los griegos dejaron de ser intratables, domesticados como irracionales con desarrollo decimal ilimitado no periódico, se usaban sin pudor aunque sin conocer demasiado bien su entidad conceptual.

El lenguaje simbólico algebraico se fue gestando a través de numerosas iniciativas, llegó a tener una gran eficacia en manos de Vieta (s. XVI/XVII), cuya pauta siguió Fermat, pero alcanzó la formulación definitiva gracias a Descartes, que por algo era filósofo. Con este lenguaje simbólico, nuevas oportunidades de desarrollo se abrieron para el álgebra.

Naturalmente, estamos hablando del conocimiento nuevo que se genera en el pequeño grupo de sabios, cada vez más comunicados entre sí, pero los caminos y las velocidades de difusión de este conocimiento hacia capas sociales más amplias requerirían un estudio específico.

La geometría analítica.

En el ambiente descrito antes, pasado el primer tercio del siglo XVII, René Descartes (1596-1650) y Pierre de Fermat (1601-1665) identificaron una buena cantidad de curvas algebraicas, aquellas cuya definición como lugar geométrico se puede escribir de la forma $F(x, y) = 0$, siendo F un polinomio que liga a los dos números variables (x, y) que determinan la posición de cada punto de la curva. Así, Fermat demostró que las cónicas de Apolonio se corresponden exactamente con las curvas de ecuaciones cuadráticas

$$F(x, y) = ax^2 + bxy + cy^2 + dx + ey + f = 0.$$

Por su parte, Descartes resolvió el llamado problema de Pappus, que pedía determinar una curva como lugar geométrico a partir de ciertas rectas, algo que nadie antes supo hacer. La solución consistió en dar las ecuaciones de la curva, que podría ser de grado arbitrario según el número de rectas dadas; a partir de las ecuaciones, el álgebra se encargaría de extraer las propiedades de la curva.

Con el tiempo, esta geometría que usa el álgebra recibió el nombre de *geometría analítica*.

El matemático hispano-argentino Julio Rey Pastor (Logroño 1888 – Buenos Aires 1962) escribió en un ensayo de 1919 que los “camino cardinales” por lo que la matemática “amplía sin cesar patrimonio” son la generalización, la especialización, la correlación y, como camino principal, “la anastomosis de problemas”, porque “los progresos esenciales corresponden a la feliz cópula de dos conceptos que parecían lejanos e independientes, y al librarlos de su envoltura formal descubren tal identidad de substancia, que se anastomosan y hasta se funden en uno solo. La Geometría analítica de Cartesio, origen de la Matemática moderna, realizó brillantemente esta anastomosis del Álgebra con la Geometría.” Con ella se trataron viejos problemas de un modo nuevo y se abordaron otros hasta entonces imposibles.

Descartes dio forma matemática a la idea vaga imperante de recta tangente a una curva algebraica, identificando la tangencia con la existencia de una raíz doble de la ecuación que resulta al cortar la recta con la curva. También introdujo el método de los coeficientes indeterminados para resolver los problemas de Apolonio sobre las normales a las cónicas y extender el tratamiento a otras curvas. Pero los métodos cartesianos tuvieron sus limitaciones, escapaban a su análisis numerosas “curvas mecánicas” a las que no se podía dar una ecuación algebraica. Nuevas funciones tenían que aflorar y nuevos métodos generales válidos para todas ellas. Un compatriota contemporáneo, De Beaune, propuso a Descartes determinar una curva conociendo datos acerca de sus tangentes; este tipo de problemas eran todavía inaccesibles, faltaba el cálculo infinitesimal.

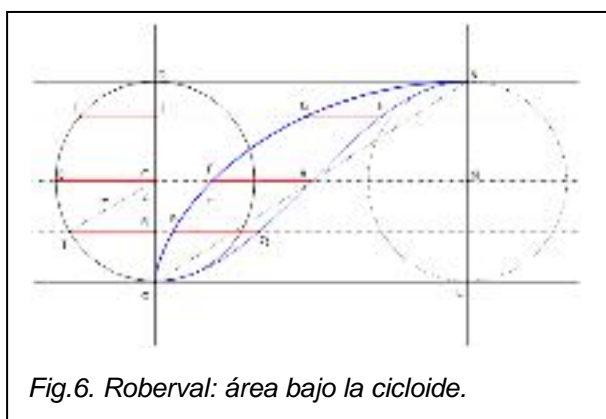


Fig.6. Roberval: área bajo la cicloide.

Últimos precursores.

Una curva mecánica famosa fue la cicloide, descrita por un punto de la circunferencia que rueda sobre una recta. Roberval, Descartes y Huygens se ocuparon de estudiarla.

El primero de ellos determinó mediante composición de movimientos la tangente en cada punto y, siguiendo la técnica

de “indivisibles” de Cavalieri, demostró que el área bajo un arco de cicloide es tres veces el área del círculo rodante que la engendra. En la Fig. 6, el área entre la cicloide $OPFGN$ y la curva compañera $OQAHN$ es igual a medio círculo generador, porque ambas figuras tienen las mismas líneas: $TR = PQ$, $BC = FA$, $IJ = GH$, etc. El área del triángulo rectángulo OLN , de base media circunferencia y altura el diámetro, es igual a la del círculo, y cambiando la hipotenusa ON por la curva compañera el área es la misma por compensación. Luego el área bajo media cicloide es vez y media la del círculo.

Fermat abordó desde la geometría analítica el cálculo con infinitesimales de Arquímedes, logró repetir con ecuaciones el cálculo del área del segmento de parábola y extender el método a otras curvas. Calculó un área complementaria mediante la aproximación por defecto y por exceso en rectángulos “muy pequeños”. Al hacerlo así, pudo extender el cálculo con la parábola ordinaria $y = x^2$ a la parábola generalizada $y = x^k$ con k un entero mayor que 2 (incluso una fracción). Para calcular el área que limita esta curva con el eje de abscisas y las ordenadas $x = 0, a$, dividió el segmento base en n partes iguales para obtener rectángulos por exceso y por defecto cuya suma acota el área que busca,

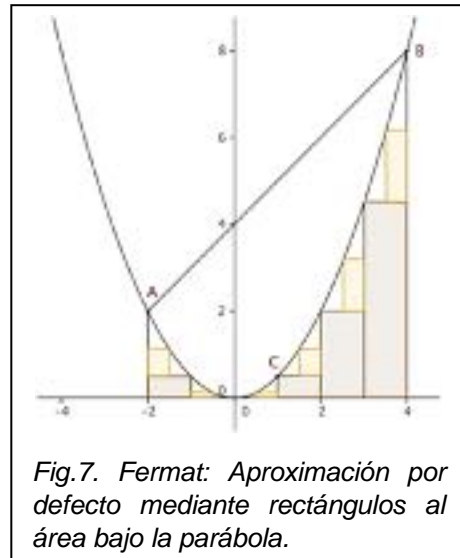


Fig.7. Fermat: Aproximación por defecto mediante rectángulos al área bajo la parábola.

$$\frac{s(n)}{n^{k+1}} < \frac{1}{k+1} < \frac{S(n)}{n^{k+1}}$$

mediante unas sumas más generales que las de Arquímedes:

$$S(n) = 1^k + 2^k + 3^k + \dots + n^k, \quad s(n) = S(n-1).$$

Estas fórmulas, deducidas del triángulo aritmético mediante inducción, eran conocidas por otros contemporáneos como Pascal, quien tenía un contacto epistolar frecuente con Fermat. Al final de los cálculos, haciendo n cada vez más grande y los rectángulos a sumar “infinitesimales”, Fermat obtiene para el área bajo la parábola de orden k el valor $\frac{a^{k+1}}{k+1}$. También aplica estos métodos a las hipérbolas generales $y = \frac{1}{x^k}$, pero esta vez dividiendo la abscisa en intervalos en progresión geométrica.

De la lectura de Apolonio y Arquímedes, obtuvo también Fermat un modo muy general de plantear el problema de la tangente vinculándolo a cuestiones de máximos y mínimos. Si una curva tiene ecuación $y = f(x)$ referida a un eje de abscisas OX , para conocer la tangente a la curva en uno de sus puntos P que se proyecta en P' sobre el eje de abscisas, Fermat busca otro punto T en el eje de abscisas tal que PT sea la tangente; lo hace calculando la longitud t de la *subtangente* $P'T$ de este modo: Del punto P' de abscisa x se desplaza a un punto “muy próximo” Q' cuya abscisa será $x + e$, siendo e un

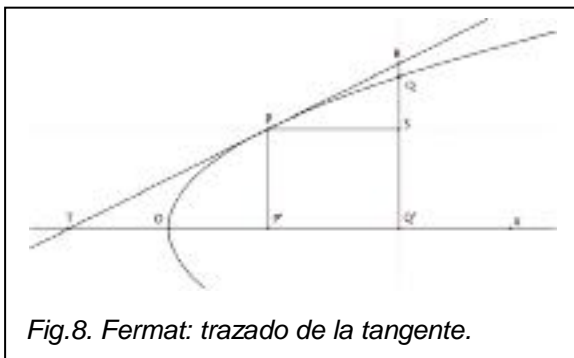


Fig.8. Fermat: trazado de la tangente.

incremento “muy pequeño”, un “infinitesimal”. Q' será la proyección de un punto Q de la curva, con ordenada $f(x + e)$ y también de un punto R sobre la tangente en P . Ya que el incremento e es “muy pequeño”, Fermat trata “como iguales” los puntos Q, R , es decir, son “como iguales” el arco de curva PQ y el segmento de tangente PR ; así que

concluye por semejanza de triángulos $f(x + e)/f(x) = (t + e)/t$, es decir,

$$tf(x + e) = (t + e)f(x).$$

De esta relación elimina el incremento e , en los casos particulares que considera, realizando un “álgebra de infinitesimales” peculiar, en la que e se comporta en las operaciones como un número $e \neq 0$ pero que finalmente se anula. Por ejemplo, para la parábola $y = x^2$:

$$t(x^2 + 2ex + e^2) = (t + e)x^2, \quad 2tx + te = x^2, \quad 2tx = x^2, \quad t = \frac{x}{2}.$$

Además de los matemáticos europeos continentales mencionados, también británicos como Gregory, Wallis o Barrow, hicieron avances similares en el cálculo de tangentes, áreas, volúmenes y también longitudes de curvas. Barrow resaltó el *cociente incremental* que da la pendiente de la tangente, con la notación anterior es

$$D = \frac{f(x + e) - f(x)}{e}$$

Se trata de una de las expresiones clave del cálculo infinitesimal, debía tener un valor explícito aunque su expresión dependiera del incremento e infinitesimal arbitrario. Calculando a la manera de Fermat, para la parábola $y = x^2$ resulta $D = 2x + e = 2x$. Desprovisto del incremento infinitesimal e , este cociente incremental es lo que ahora llamamos la *derivada* $y' = 2x$. Análogamente, la derivada de $y = x^n$ es $y' = nx^{n-1}$.

Newton y Leibniz.

Llegado el último tercio del siglo XVII, del grupo de matemáticos herederos de Arquímedes que calculaban con los infinitesimales, emergieron dos más penetrantes que fundaron al fin el cálculo infinitesimal: Newton y Leibniz. En el escritorio de estos genios, el cálculo infinitesimal se originó al concebir el cálculo de tangentes y de áreas de manera unitaria. El método de calcular tangentes fue llamado *cálculo diferencial* y el de calcular longitudes, áreas y volúmenes *cálculo integral*, estas son las dos partes que se anastomosan para formar el cálculo infinitesimal.

El cálculo diferencial recibe este nombre porque Leibniz llamó *diferenciales* (diferencias pequeñas) a los pequeños incrementos de una variable x y los designó dx (otra vez un filósofo acertando con el signo conveniente). Así, cuando en la relación $y = f(x)$ la cantidad variable x recibe un incremento diferencial pasando a ser $x + dx$, la variable dependiente y pasa de $f(x)$ a $f(x + dx)$, siendo

$$\frac{dy}{dx} = \frac{f(x+dx)-f(x)}{dx}.$$

Se dice que la derivada $y' = f'(x)$ “es igual a” $\frac{dy}{dx}$ una vez que este cociente incremental se ha calculado como un número determinado en función solo de x , eliminando las diferenciales, lo que Newton y Leibniz hicieron en un principio usando la misma “álgebra de infinitesimales” que Fermat.

Newton pensaba que las variables eran siempre dependientes del tiempo y al producirse un cambio infinitesimal en el tiempo cambiaban de un modo lineal, apareciendo la derivada como coeficiente: se pasaban de $x(t)$ en el instante t a ser un “momento” después

$$x(t + o) = x(t) + \dot{x}(t)o.$$

Newton denotaba o al pequeño incremento del tiempo t y escribía \dot{x} para la derivada, a la que llamaba la *fluxión* de la cantidad x que *fluye*. Además, imaginaba la tangente como un secante que cortaba a la curva en dos puntos muy próximos que se confundían en uno.

En el cálculo de áreas, de nuevo el filósofo Leibniz acertó con el signo adecuado: expresó el área limitada por la curva $y = f(x)$, el eje de abscisas y las ordenadas $x = a, b$, de la forma $\int_a^b f dx$, donde el símbolo \int representa la inicial “s” estilizada de la palabra “suma”, recordando la visión del área como “suma de líneas” que tuviera Cavalieri. Con esta notación, el resultado de Fermat comentado antes se escribe

$$\int_0^a x^k dx = \frac{a^{k+1}}{k+1}.$$

Donde Newton y Leibniz marcaron la diferencia con sus predecesores y contemporáneos fue en producir la anastomosis entre el cálculo diferencial y el integral, usando la nueva visión del cálculo infinitesimal para resolver problemas hasta entonces sin solución. La esencia de la anastomosis radica en la consideración de la integración (áreas) como operación inversa de la derivación (tangentes) a través de lo que se dio en llamar el *teorema fundamental del cálculo*, un resultado que conocía Barrow, maestro de Newton, pero le pasó desapercibida su crucial importancia, siendo para él uno más en el bosque de sus teoremas geométricos.

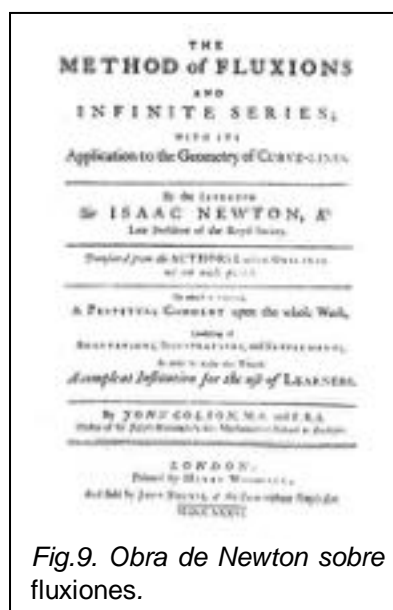


Fig.9. Obra de Newton sobre fluxiones.

Viendo el área bajo la curva $y = f(x)$ como una magnitud variable

$$A(x) = \int_a^x f dx,$$

Newton y Leibniz consideraron como una cuestión de principio, que cada uno argumentó a su manera, que la derivada de esta función es la función inicial: $A'(x) = f(x)$. Como al derivar la suma de constantes no cuenta, cualquier función $F(x)$ que verifique $F'(x) = f(x)$ (llamada *primitiva* de $f(x)$ y denotada $\int f dx$) servirá para calcular la integral mediante el teorema fundamental del cálculo:

$$\int_a^b f dx = F(b) - F(a).$$

Para cerrar el proceso de invención del cálculo infinitesimal falta saber calcular derivadas y primitivas de funciones. Esto estaba claro con los polinomios, basta saber hacerlo con el caso simple $y = x^k$, ya conocido por Fermat y otros incluso con exponentes fraccionarios.

Pero en tiempos de Newton y Leibniz iban apareciendo otras relaciones funcionales, no solo las fracciones y las raíces con polinomios, también las vinculadas al trabajo que los astrónomos hacían para la navegación, los senos y cosenos junto con otras razones trigonométricas y sus inversas. A ellas se añadieron a lo largo del siglo XVII los logaritmos, creados por el astrónomo Napier y el relojero Briggs a principios del siglo y más tarde relacionados con las áreas bajo la hipérbola $y = \frac{1}{x}$ por matemáticos como Saint Vicent y Sarasa. Otro matemático muy vinculado a los asuntos de la navegación y los mapas, Mercator, trabajando al estilo de los indivisibles de Cavalieri, encontró en los inicios del último tercio del siglo una variante funcional de las sumas infinitas numéricas, dando para el cálculo de logaritmos el desarrollo en serie de potencias

$$\log(1 + x) = x - \frac{x^2}{2} + \frac{x^3}{3} - \frac{x^4}{4} + \dots,$$

que permitía mayor facilidad y precisión en la obtención de tablas. A esta suma infinita funcional, se sumaron otras con el nombre de *series de potencias*, como esta que resulta del simple algoritmo de división de polinomios:

$$\frac{1}{1 - x} = 1 + x + x^2 + x^3 + \dots,$$

cuyos sumandos van decreciendo cuando $x < 1$. Vimos los casos $x = \frac{1}{2}, \frac{1}{4}$, unas páginas atrás. En estos asuntos trabajaba también Wallis, otro de los maestros de Newton, gran experto en sumas infinitas.

Este es el recurso que en manos de los genios Newton y Leibniz cerró la bóveda del cálculo infinitesimal, consiguieron desarrollar en series de potencias las funciones elementales entonces conocidas y las consideraron como

“polinomios con puntos suspensivos”, derivando e integrando término a término en dichos desarrollos. Así, resultará

$$\log'(1+x) = \frac{1}{1+x}, \quad \int \frac{1}{1+x} dx = \log(1+x).$$

En la obtención de los desarrollos en series de las funciones tuvo un papel muy destacado la extensión que Newton hizo de la conocida fórmula del binomio para expresar las funciones $(1+x)^k$ con exponente entero negativo o fraccionario como una serie de potencias con infinitos términos.

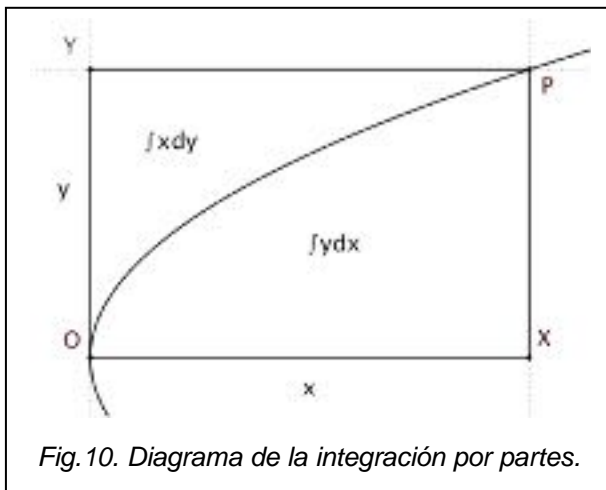


Fig.10. Diagrama de la integración por partes.

Enseguida se dieron las reglas para derivar productos y cocientes, se crearon tablas de primitivas y se formuló la integración por partes,

$$xy = \int x dy + \int y dx,$$

etc. El cálculo infinitesimal se puso en marcha.

Para verificar su potencia y dejar claro que se trataba de un gran paso de la humanidad, cada uno de los dos genios fundadores resolvió problemas hasta entonces imposibles. Leibniz resolvió el problema que De Beaune había planteado a Descartes, lo hizo mediante una ecuación diferencial cuya solución encontró usando logaritmos. Como parte de sus investigaciones sobre el movimiento, Newton completó el estudio de la curvatura de las curvas que había iniciado Huygens.

La gran obra de Newton, uno de los logros más imponentes del pensamiento humano, fue formular las leyes generales del movimiento de los cuerpos, estudio iniciado por Galileo, y deducir de ellas las leyes de Kepler inducidas a partir de datos observacionales. Para la exposición de su dinámica, Newton prefirió seguir el método sintético de la geometría de Euclides y no el analítico, lo que dificultó la comprensión de su obra y la de sus seguidores británicos, como Maclaurin o Taylor. Pero matemáticos posteriores de la Europa continental desarrollaron y completaron la teoría newtoniana haciendo uso del cálculo infinitesimal analítico, que es el modo natural de expresión matemática de dicha teoría.

El cálculo alienta el progreso y se reformula.

La geometría analítica y el cálculo infinitesimal empezaron inmediatamente a dar frutos en la matemática pura (resolución de problemas matemáticos) y en las matemáticas mixtas (aplicadas a la mecánica y otras ciencias). A finales del

siglo XVII y principios del XVIII ya había una amplia cantidad de problemas importantes resueltos con los nuevos métodos, principalmente por los fundadores y los hermanos suizos Johann y Jakob Bernoulli, seguidores de Leibniz.

A lo largo del siglo XVIII la geometría analítica del plano y del espacio quedó consolidada, pero el cálculo infinitesimal todavía no. Aparecieron libros que ayudaban a estudiarlo: el del Marqués de L'Hospital, fruto de las lecciones recibidas de Johann Bernoulli, y el de la italiana Caetana Agnesi, ambos muy ligados al estudio de curvas. Siguiendo la onda continental, en la segunda parte del siglo llegaron los notables libros de Euler y Lagrange

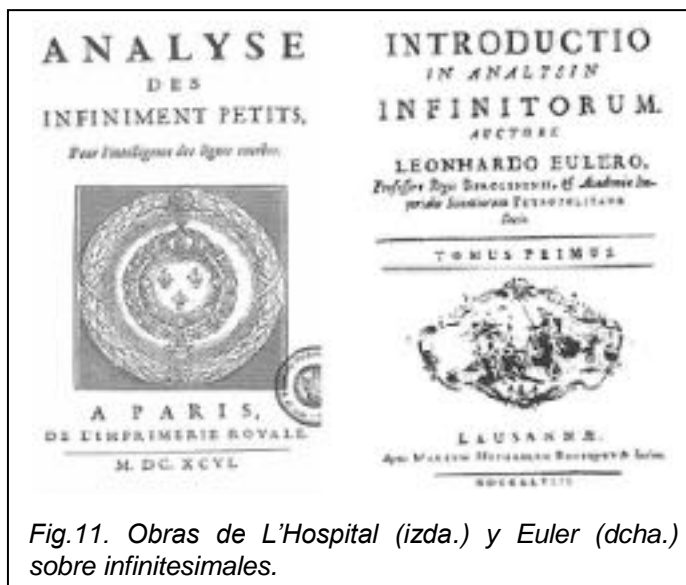


Fig.11. Obras de L'Hospital (izda.) y Euler (dcha.) sobre infinitesimales.

planteando el cálculo como una teoría de las funciones autónoma, dejando las cuestiones sobre curvas y superficies para las aplicaciones. Por otra parte, ambos autores expresaron la dinámica de Newton con la nueva herramienta matemática, lo que alcanzó forma acabada en la *Mecánica analítica* de Lagrange (1788), que refrendó el valor indudable del cálculo infinitesimal.

Pero la matemática pura exige también, desde los *Elementos* de Euclides, que las justificaciones de los métodos empleados estén bien consolidadas con lógica deductiva y en esto la situación inicial del cálculo infinitesimal fue deficiente: triunfaba como método de invención y aplicación, pero adolecía de una justificación suficiente. En la *Enciclopedia* francesa de Diderot y D'Alembert aparecen las cantidades infinitesimales o diferenciales como parte de los registros matemáticos, pero también en el apartado de metafísica, donde se discute si son cantidades reales o meras ficciones sometidas a un álgebra específica y capaces de producir explicaciones de la realidad. En este aspecto los infinitesimales eran como el éter o el vacío.

D'Alembert aconsejaba a sus estudiantes que no hicieran demasiadas preguntas sobre el fundamento del álgebra de los infinitesimales, pero que se afanaran en aplicarlos con éxito a los más diversos problemas prácticos cuya solución impulsaba el progreso. Sin embargo, él mismo se ocupó de salvar el escollo metafísico, fue un adelantado en la teoría de los límites que usamos hoy, en la que radicaba, decía, "la verdadera metafísica del cálculo infinitesimal".

Ya en el siglo XIX, su compatriota Cauchy (1821) hizo desaparecer los infinitesimales sustituyéndolos por la teoría de límites que se estudia hoy, en la que la derivada se expresa como un límite del cociente incremental:

$$f'(x) = \lim_{e \rightarrow 0} \frac{f(x+e) - f(x)}{e}$$

También las integrales se definen desde Cauchy como límites de sumas de áreas por exceso y por defecto. No obstante, cuando se usa el cálculo con fines prácticos, siguen siendo valiosas las intuiciones de los infinitesimales.

Hacia los años treinta, Fourier dio un nuevo impulso heurístico al cálculo con su teoría del calor, en la que usaba con inusitada libertad desarrollos de las funciones en series infinitas de senos y cosenos, cuyos coeficientes calculaba mediante integrales. Las series trigonométricas de Fourier tuvieron un enorme éxito aplicado y necesitaron una costosa adaptación al rigor conceptual, al que se aplicaron los matemáticos europeos que dieron forma acabada al cálculo infinitesimal cuando ya recibía nombres como *teoría de funciones* o *análisis matemático*. Para ello hubo que esperar hasta el último tercio del siglo XIX, cuando se fundamentó con rigor aritmetizado (Cantor, Dedekind, Weierstrass) el número real que modela el continuo.

Mediante el cálculo infinitesimal, Gauss completó el estudio de la curvatura de las superficies (1827) y a partir de su trabajo concibió Riemann (1854) la noción abstracta de un espacio matemático curvo y multidimensional con diferentes tipos de métrica, dejando para los físicos la tarea de determinar el modelo matemático que corresponde a la realidad material que ellos examinan. Cuando, en los primeros años del siglo XX, Einstein formuló los principios físicos de la mecánica relativista, la exposición matemática de su teoría se realizó en un espacio-tiempo con métrica de Riemann.

Bibliografía, lecturas de ampliación (solo en lengua española):

- (1) Boyer, C. B., *Historia de la matemática*, Alianza Ed., Madrid, 1991.
- (2) Durán, A. J., *Una historia, con personajes, de los conceptos del cálculo*, Alianza Ed., Madrid, 1996.
- (3) Kline, M., *El pensamiento matemático de la antigüedad a nuestros días*, Alianza Ed., Madrid, 1992.
- (4) González Urbaneja, P. M., *Las raíces del cálculo infinitesimal en el siglo XVII*, Alianza Ed., Madrid, 1992.
- (5) Rey Pastor, J., *Curso de cálculo infinitesimal*, Madrid, 1962. Con un apéndice sobre la evolución y los precursores del cálculo infinitesimal.
- (6) Rey Pastor, J., Babini, J., *Historia de la matemática*, 2 vols., Gedisa, Barcelona, 1984.
- (7) Thompson, S. P., *El cálculo infinitesimal al alcance de todos*, Madrid, 1932.

Luis Español González

Profesor Titular de Geometría y Topología.

Historiador de las Matemáticas.

Departamento de Matemáticas y Computación.

Universidad de La Rioja.

Capítulo 7

La Electroquímica: su historia y repercusión en distintos campos.

Introducción.

La importancia de la Electroquímica está justificada solo con revisar los muchos campos del saber en los que está implicada: Basta con repasar los programas de cualquier congreso relacionado con Electroquímica por ejemplo la reunión del 10 al 11 de julio de 2017 de Berlín incluía [1]:

Principios de Electroquímica

Electroquímica Teórica y Computacional

Química Electroanalítica y Ciencia de Sensores

Energía Electroquímica Conversión y Almacenaje

Bioelectroquímica

Ciencia electroquímica de superficies

Baterías y fuentes de energía

Electroquímica Orgánica y Organometálica

Corrosión, Ciencia y Tecnología

Electroquímica Física y Analítica, Electrocatalisis y Foelectroquímica

Aplicaciones de Electroquímica:

- Galvanoplastia y Recubrimientos

- Orgánica y Bioelectroquímica

- Ingeniería Electroquímica

- Nanoestructuras y Dispositivos de Carbono

Sin duda, la Electroquímica es una ciencia de muchos tentáculos con presencia en un gran número de ciencias y, por otra parte, añadiríamos que para los profesores es un tema ideal para seguir la enseñanza STEAM (Science, Technology, Engineering, Arts, Mathematics).

Consideramos que Jöns Jacob Berzelius (1779-1848) se quedó corto cuando en su "Tratado de la química mineral, vegetal y animal", con gran intuición, decía: "*El descubrimiento de la electricidad desarrollada por medio de*

contacto y los efectos a que ha dado lugar, es quizá, atendiendo a la influencia que ha ejercido sobre las ramas de la física, uno de los más importantes que ha podido hacer el entendimiento humano. Cuando se reflexiona al estado que ha llegado la **filosofía química** después de esta época, y se recuerda que hace treinta y cinco años ni siquiera se sospechaba, debemos felicitarnos de haber nacido en un siglo en el que el género humano marcha con más rapidez que en ninguno otro tiempo hacia su perfección intelectual" [2].

Descubrimiento de la pila. [3,4,5,6]

Podemos considerar que la historia comienza cuando, el médico y físico italiano, Luigi Galvani (1737-1798) se interesó por la fisiología y su relación con la electricidad porque al diseccionar en 1780 una pata de una rana tocó el bisturí con un gancho de bronce del que colgaba la rana y ésta se contrajo de forma espontánea. Él lo achacó a un fenómeno eléctrico, siguió haciendo múltiples experimentos para comprobar si se trataba de una electricidad de origen animal y los resultados los dio a conocer en *De viribus electricitatis in motu musculari* [7]. Este descubrimiento llevó a su amigo, el físico italiano

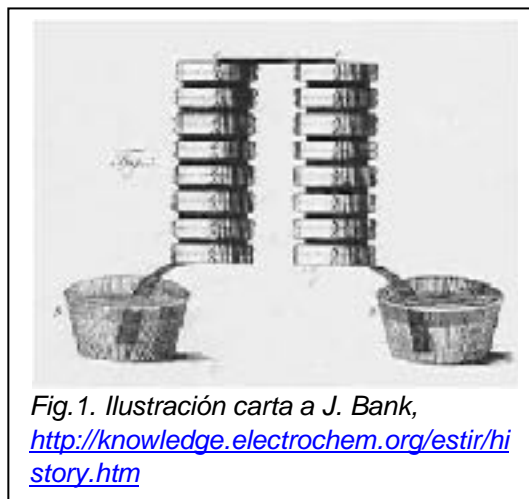


Fig.1. Ilustración carta a J. Bank, <http://knowledge.electrochem.org/estir/history.htm>



Fig.2. A. Ganot Física, 1868, 478

Alessandro Giuseppe Antonio Anastasio Volta (1745-1827), a estudiar el fenómeno que pronto estuvo en contra de la idea de Galvani sobre la necesidad de la presencia del animal y demostró que únicamente con dos metales y un medio húmedo entre ellos se producía el mismo fenómeno hecho que describe (en francés) en su famosa carta de marzo de 1800 dirigida a Joseph Banks, presidente de la Royal Society, diciendo que consta de discos de metal y discos de cartón húmedos, la mejor opción de metales: "cobre, o mejor plata" y "láminas de estaño, o lo que es mucho mejor, zinc", agua salada que impregna "algún material esponjoso". Puntualiza que es importante que los discos metálicos estén muy limpios y secos. En resumen: son plata discos húmedos cinc, otra vez plata discos húmedos cinc y así hasta 20 de pares de metales. Añade que buscó formas de alargar mucho la columna y en la carta dibuja una serie de cuatro columnas idénticas conectadas con lo cual llega a 80 voltios.

Volta se dedicó a investigar la acción de la pila sobre metales, animales, etc. y en ese contexto se descubrió la corriente eléctrica. No solo los científicos también los políticos se interesaron por este descubrimiento y Volta en 1801 hizo una demostración ante Napoleón Como consecuencia de lo impresionado que quedó Napoleón dotó

a los científicos franceses para que construyeran la pila más grande del mundo [8,9]. Las pilas que llegaron a ser monumentales porque las de columnas se sustituyeron por pilas de vasos que ocupaban habitaciones enteras, con el descubrimiento por Faraday en 1831 de las corrientes inducidas que llevaron a los generadores eléctricos de inducción perdieron todo interés y solo se siguió utilizando la pila seca descubierta en 1868 por Georges Leclanche (1839-1882) y que fue perfeccionando hasta conseguir la pasta despolarizante que le permitía prescindir del vaso poroso.

La pila de Leclanche sufrió varias transformaciones hasta que Carl Gassner (1855-1942) en 1887 consiguió una pila fácil de transportar que aún se utiliza. En 1942 en el ataque a Pearl Harbor las pilas secas con el calor explotaban y no le funcionaban los instrumentos de transmisión por lo que le pidieron a Samuel Ruben (1900-1988), que tenía fama de ser “un manitas”, que buscara una solución y fabricó la pila de botón con óxido de mercurio como despolarizante. Sobre 1970 se fue encontrado mercurio en los lugares más impensados dentro de la cadena trófica hasta que se dieron cuenta de que procedía de las pilas alcalinas que se abandonaban en cualquier sitio y por eso se empezó a investigar en una posible pila de litio porque al tener el litio potencial -3.0 V, el doble de 1,5 de las de Leclanche se podrían acoplar a los mismos aparatos. Figuran como promotores o padres de la batería de litio sobre 1970 Stanley Whittingham (1941-) y 1980 John Bannister Goodenough (1922-) son recargables y de las que en este momento hay varias patentes y en las que se sigue trabajando desde la industria de automóviles, ordenadores, móviles, etc. La primera batería recargable fue el acumulador de plomo de Gastón Planté (1834-1889) de 1859, que a pesar de sus problemas: de peso por el plomo y de corrosión por el sulfúrico, se ha seguido usando hasta hace pocos años que como hemos indicado se ha sustituido por baterías de litio.

El origen de las pilas de combustión está en una carta del 22 de octubre de 1842, que William Robert Grove (1811–1896) dirige a Faraday donde le explica que ha completado una curiosa pila voltaica que supone que le gustará ver, está formada de tubos alternando de hidrógeno y oxígeno a través de los cuales pasa un muelle de platino, cada par está sumergido en agua acidulada de forma que el líquido toca el extremo del muelle de

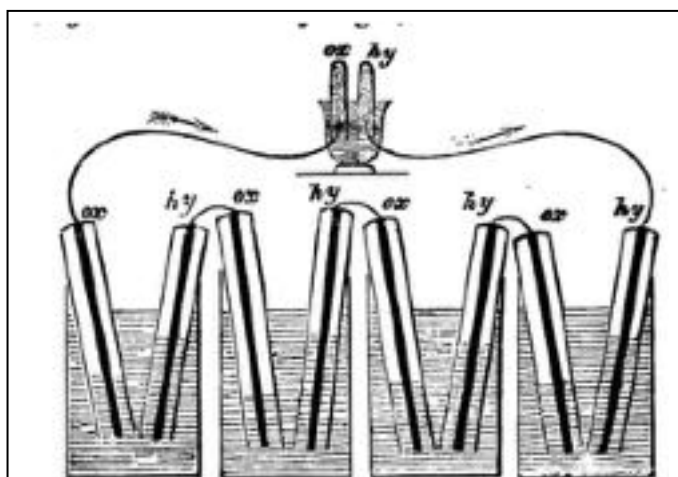


Fig.3. Pila de Grove. Carta a Faraday *Phil. Mag.* Ser. III, vol XXI,p.418

platino platinizado y con ella ha conseguido electrolizar el yoduro potásico y el ha sufrido una fuerte descarga. En 1842 Grove utilizó dicha pila para iluminar con lámparas incandescentes la sala de lectura de la Real Sociedad [4,10].

La pila de combustión teóricamente es el generador ideal de corriente eléctrica porque sería el más rentable ya que los electrones pasan directamente del combustible al comburente dando lugar a la corriente y sin necesidad de producirse primero una energía térmica que caliente agua y que el vapor mueva el generador electromagnético, lo que supone una serie de transformaciones de energía con las correspondientes pérdidas. En la actualidad existen diversas variedades pero su uso es todavía muy limitado y se sigue investigando.

Descubrimiento de la electrólisis. [3,4,5,6]

William Nicholson (1753-1815) y Anthony Carlisle (1768-1840) descubren la electrólisis al darse cuenta que al pasar la corriente por el agua se desprenden gases, hecho que observó Nicholson en una gota de agua que había colocado para hacer contacto en la pila, lo que les llevó a estudiarlo en profundidad. Pero son Humphry Davy (1778-1829), Berzelius y Michel Faraday (1791-1867) los que dan un gran impulso a la utilización y conocimiento de este fenómeno, los dos primeros porque por electrólisis obtienen o descubren la mayoría de los metales alcalinos y alcalinotérreos (Na, K, Mg, Ca, Sr, Ba y B), y el segundo porque cuantifica el fenómeno y además es el padre de la nomenclatura de la electrólisis.

Berzelius junto con Wilhelm Hisinger (1766 –1852) físico y químico sueco sometiendo las sustancias más diversas a la acción de la pila de Volta e intentando clasificar los productos se dieron cuenta de que todas las sustancias inorgánicas están compuestas de una parte positiva y otra negativa lo que supuso un gran avance en la ciencia química investigaciones que publicaron en 1803 en su trabajo titulado “*Experimentos sobre el efecto del electricidad en las sales y en algunas de sus bases*”, aunque fue más conocido a partir de 1807 cuando se publicó en *Annalen der Physik* [11,12].



Fig.4. *Ann.Phys*, 1807, XXVII.,270
<https://gallica.bnf.fr/ark:/12148/cb34462944f/date>

Johann Wilhelm Ritter (1776 - 1810) físico y filósofo alemán, comprobó que la escala de potenciales de Volta coincidía con la de afinidades químicas. Davy y Berzelius constataron que los fenómenos que se producían en la electrólisis y en las pilas eran los mismos. De esta forma se sentaron las bases de la electricidad debida a la acción química, que es la **electroquímica**, palabra acuñada por el físico francés Auguste Arthur de la Rive (1801-1873).

Primero William Hyde Wollaston físico y químico británico (1776 –1828) que perfeccionó la pila y después Davy y Berzelius vieron que la electrólisis se producía lo mismo con la electricidad galvánica que con la producida por frotamiento y llegaron a la conclusión de que ambas eras iguales.

Davy en 1806 en la conferencia “bakerian” de la Real Sociedad, uno de los acontecimientos científicos más importantes, afirmó la identidad de las reacciones de descomposición de las sustancias y las que daban origen a la electricidad de contacto, lo que suponía que las dos tenían que ver con la afinidad química por lo que consideraba a los compuestos químicos como un equilibrio entre las energías eléctricas naturales de los cuerpos y por eso cualquier tipo de electricidad podía romper ese equilibrio descomponiendo las sustancias. Al final de su conferencia hizo notar la importancia que tendría ese procedimiento para la obtención industrial de diferentes sustancias.

De dicha conferencia hizo grandes elogios Berzelius lo que contribuyó a aumentar la fama de Davy y el Instituto de Francia le concedió el premio que unos años antes había creado Napoleón para premiar a Volta. Francia e Inglaterra en ese momento estaban en guerra y la opinión general era que no debía de aceptar aquella distinción pero Davy contestó: *“Aunque los países estén en guerra los sabios no están”*.

Davy se dio cuenta de que cuando hacia la electrólisis de las disoluciones de sales, bases, etc., siempre se descomponía el agua, para eliminar el agua intentó descomponer las sustancias sólidas con pilas de Volta muy potentes sin ningún éxito porque no conducían la electricidad por lo que optó por fundirlas. Comenzó por fundir potasa en una cuchara de platino, y consiguió una masa roja fluida de potasa quemando alcohol con oxígeno. Al someterla a la acción de la pila en el hilo negativo se produjo una llama muy intensa. Repitió el experimento varias veces y en distintas condiciones con potasa totalmente pura y seca, en una cuchara aislada de forma que la pila se conectara a la sustancia fundida y comprobó que aparecían unos glóbulos metálicos tan brillantes como si fueran de mercurio. Algunos se quemaban rápidamente y otros se recubrían de una capa blanca. Estaba claro que había conseguido aislar el potasio el 6 de octubre de 1807 y unos días más tarde aisló el sodio. Los trabajos de Berzelius y, su experto ayudante, Hisinger completaron el descubrimiento y la obtención de la mayoría de los metales activos.

La antorcha para seguir avanzando en Electroquímica la recogió Faraday, conocido como el Newton de la electricidad. Por una parte entre 1830 y 1833 primero con la colaboración de su médico Dr. Whitlock Nicholl (1786-1838) y después con la ayuda de William Whewell (1794-1866), del Trinity College de Cambridge, acuñó la terminología de la electrólisis intentando que

fuera una descripción correcta de lo que allí sucedía, cómo se transmitía la electricidad y que procesos se producían en la celda. No estaba de acuerdo que la acción de la pila fuera como decía en 1805 Johann Dietrich von Grothuss (1785-1822) una acción que iba pasando de unas moléculas a otras, opinión que también defendía Davy, sino a una partes que se desplazaban en la disolución por eso los llamó iones (viajeros) apoyando las teorías de William Robert Grove (1811-1896) que suponía que las moléculas tenían que estar rotas antes de la electrólisis porque si no se necesitaría más energía.

Con ayuda de Nicholl introdujo los términos: *electrolisis* (de *electro-* y *-lisis*; desatar, romper una sustancia por la electricidad), *electrolito* (de *electro-* y el griego λυτός, *lytós*, “soluble”) y *electrolizar* (de *electrólisis*) La terminación “-lisis” procede del griego λύσις “desatar”, y el termino “*electro-*” del latín *electrum*, el cual a su vez procede del griego ἤλεκτρον, “ambar” Cuando Faraday trata de nombrar los conductores que están en el interior de la celda a ninguno de los dos les acaban de convencer lo que se les ocurre, por lo que deciden recurrir a William Whewell que les recomienda, después de largas discusiones, que se queden con las palabras *ánodo* y *cátodo*, del término “*odo*” (de odos, οἶος; camino o vía), cuyos significados serían “camino hacia arriba” y “camino hacia abajo”, respectivamente [13].

Faraday se dio cuenta que para producir la electrólisis eran suficientes tensiones muy pequeñas capaces de vencer las afinidades químicas y haciendo estudios cuantitativos llegó a las leyes de la electrólisis, que podía haber descubierto Berzelius, treinta años antes, si no hubiera confundido intensidad con cantidad de electricidad. Las leyes las enunció en 1832 diciendo:

“La energía química lo mismo que la magnética son directamente proporcionales a la cantidad de electricidad que pasa”

“Los pesos equivalentes de las sustancias son simplemente aquellas cantidades de ellas que contienen iguales cantidades de electricidad”

“Siendo la electricidad la que determina el número de equivalentes porque es la que determina la fuerza de la combinación. O si adoptamos otra fraseología de la teoría atómica de los átomos de los cuerpos que son equivalentes en una acción química ordinaria tiene iguales cantidades de electricidad asociadas naturalmente a ellos. Pero debo confesar que estoy muy celoso de usar el término átomo, porque es muy difícil tener una idea clara de su naturaleza especialmente si se consideran cuerpos compuestos”

Según afirmó años más tarde Louis-Victor Pierre Raymond de Broglie (1892-1987) las leyes de Faraday podían hacer pensar en la estructura discontinua de la electricidad y en el átomo de electricidad, fue Hermann Ludwig Ferdinand von Helmholtz (1821–1894) el primero que en 1891 afirmó que cada átomo o grupo de átomos de la electrólisis transportaban átomos de electricidad y para esos átomos George Johnstone Stoney (1826-1911) propuso el nombre de electrón.

Faraday fue quien utilizó por primera vez la medida de los volúmenes de los gases desprendidos en la descomposición del agua para cuantificar la acción de la pila de Volta y al dispositivo lo llamó volta-electrómetro pero a sugerencia de John Frederic Daniell (1790-1845) cambio a voltámetro, era más corto y sonaba bien. Indicaba, que para hacer medidas, no había sustancia más adecuada que el agua cuando se convierte en buena conductora añadiéndole un ácido o una sal por lo fácil que resulta recoger y cuantificar los gases desprendidos aunque tuvo que resolver problemas como la adsorción o combinación con los electrodos y describió hasta cinco voltámetros diferentes.

Theodore William Richards (1868 –1928) años más tarde para hacer unos cálculos más precisos utilizó una disolución de nitrato de plata y sustituyó el voltámetro por el culombímetro. Estos trabajos le permitieron hacer medidas muy precisas de pesos atómicos y de los diferentes isótopos por lo que le

dieron el Premio Nobel de Química en 1914. Sus medidas fueron la base para definir el amperio durante mucho tiempo como la intensidad de una corriente que deposita 1,118 mg de plata en un segundo.



Fig.5. Voltámetro de Hoffman
(CT/EDU/0026)

http://webs.ucm.es/info/diciex/proyectos/pie_2002-42/ct.edu.0026.html

Electroquímica iónica. [3,4,5,6]

El siglo de XIX además del siglo de oro de la electricidad fue el de la termodinámica que comenzó con el artículo de 1842 de Julius von Mayer (1814- 1878) “*Sobre las fuerzas de naturaleza inorgánica*”, con la determinación del equivalente mecánico del calor por James Prescott Joule (1818-11 1889), la generalización del principio de equivalencia de Hermann Ludwig Ferdinand von Helmholtz (1821-1894), la Teoría Cinética de Rudolf Julius Emmanuel Clausius (1822-1888) y la Mecánica Estadística de James Clerk Maxwell (1831-1879) y Josiah Willard Gibbs (1839–1903). Faraday cita las ideas de Mayer y Joule sobre conservación de la energía y está convencido de que existe un principio general que abarca todos los fenómenos naturales, por lo que esas equivalencias deben existir también en los fenómenos eléctricos.

En 1878 Jacobus Henricus van't Hoff (1852–1911, Premio Nobel de Química 1901) junto con François-Marie Raoult (1830-1901), descubren que dentro de las disoluciones que siguen las leyes de la ebulloscopia y crioscopia hay unas cuyos valores experimentales coinciden con los teóricos pero en otras se necesita un factor corrector mayor que la unidad.

Svante August Arrhenius (1859 –1927) defendió su tesis doctoral en 1887 teniendo que enfrentarse a un duro debate, incluso estuvieron a punto de suspenderlo, con un trabajo que le valió el Premio Nobel de Química de 1903 y en el que explicaba a qué se debían la anomalías que habían encontrado van't Hoff y Raoult. Arrhenius dice: *“Con objeto de explicar los fenómenos eléctricos suponemos como Clausius que algunas de las moléculas de los electrolitos están disociadas en sus iones, los cuales se mueven de forma independiente... si fuéramos capaces de calcular la fracción de moléculas de electrolito disociadas en iones podríamos calcular la presión osmótica”....”designaría las moléculas cuyos iones son independientes como activas y las restantes, cuyos iones permanecen unidos como inactivas”*. Friedrich Wilhelm Georg Kohlrausch (1840-1910) midiendo conductividades obtuvo entre 1866 y 1886 unos datos experimentales que apoyaban la teoría de Arrhenius.

El descubrimiento del electrón por Joseph John Thomson (1856- 1940 Premio Nobel de Física 1906) y la teoría iónica de Arrhenius confirmaron la diferencia entre la conducción eléctrica de los metales por el movimiento de los electrones y de los electrolitos por el movimiento de los iones y permitieron comprender, una buena parte, de los fenómenos electroquímicos. Estos descubrimientos fueron también un interesante apoyo para el avance de la Química como ciencia porque permitieron explicar las reacciones de las disoluciones acuosas de las sustancias iónicas, las teorías ácido- base, etc.

De esta forma, con Arrhenius nació la electroquímica iónica que se completaría con los trabajos de Friedrich Wilhelm Ostwald (1853-1932) premio Nobel de Química en 1909, que demostró que solo con diluciones infinitas se podría llegar a un aumento del grado de disociación. A Arrhenius, Ostwald y van't Hoff se les conocía como los pioneros. Van't Hoff fue el único que apoyó a Arrhenius desde el principio.

Todavía quedaban flecos pendientes de resolver y, en parte, encontró la solución Niels Janniksen Bjerrum (1879–1958), que en 1909 introdujo el coeficiente osmótico que permitía cuantificar las disoluciones no ideales de electrolitos, pero el problema no se solucionó totalmente hasta los descubrimientos de Peter Joseph William Debye (1884–1966, Premio Nobel de Química 1936) y Erich Armand Arthur Joseph Hückel (1896-1980) en 1923 y los de Lars Onsager (1903-1976 Premio Nobel de Química en 1968) en 1926-27. Por ese motivo la ecuación que explica los valores reales de las conductividades lleva los nombres de esos tres científicos. Las diferencias entre las disoluciones ideales y reales se explican teniendo en cuenta que los iones o las moléculas polares se unen a cualquier ión. De hecho el premio nobel de química a Debye se le concedió por sus estudios sobre los momentos dipolares.

La acción de las moléculas polares del disolvente explican la observación que en 1893 habían hecho J.J. Thomson y Walther Hermann Nernst Görbitz (1864-1941 premio Nobel de Química en 1920) sobre la influencia de la constante dieléctrica del medio en la fuerza de atracción de los iones, disminución que explico Ralph Howard Fowler (1889-1944) en 1939 aplicando cálculos de mecánica estadística.

El primero que aplicó la termodinámica a la pila de Volta fue Lord Kelvin (William Thomson, 1824-1907), diciendo que la energía producida por la reacción química debía de ser igual que el trabajo que producía, sin embargo no se dio cuenta que se podía gastar trabajo en el interior de la pila. Unos años más tarde, primero Helmholtz y después Gibbs intentaron resolver el problema de la pila aplicando el ciclo de Nicolas Léonard Sadi Carnot (1796-1832). Tomando como base todos los trabajos anteriores Nernst llegó a la ecuación que lleva su nombre y que aparece publicada por primera vez en 1889 referida a distintas disoluciones separadas por una membrana semipermeable, tema que presentó para su habilitación como profesor de la Universidad de Leipzig.

Nernst siguió trabajando con objeto de calcular la fuerza electromotriz de la pila e introdujo el concepto de presión o tensión electrolítica que medía la tendencia del metal a disolverse, es decir, cambiar de estado atómico a estado iónico, lo mismo que la tensión de vapor mide la tendencia de pasar de líquido a vapor, afirmando que para obtener la fuerza electromotriz es necesario tener en cuenta cuatro magnitudes: Dos metal-disolución una para cada electrodo, una disolución-disolución y otra metal-metal. De esta forma llegó a la ecuación que se conoce con su nombre en 1909 y por la que obtuvo el Premio Nobel de Química en 1920. Esta ecuación es incorrecta porque desprecia las dos últimas diferencias de potencial y además supone que las disoluciones son ideales y siguen las leyes de los gases perfectos.

El estudio de esta ecuación lleva a deducir que se puede escribir como si fueran dos potenciales de electrodo y llegar a ella por la diferencia entre los dos. Como no se podían medir los potenciales de forma aislada se decidió tomar el potencial del hidrógeno como término cero de la escala de potenciales. Se consideró cero la diferencia de potencial que existe entre un electrodo de platino saturado de hidrógeno y una disolución normal con respecto a los iones hidrógeno. En 1913 Joel Henry Hildebrand (1881–1983) perfeccionó la definición de electrodo de hidrógeno, especificando “*se hace introduciendo un electrodo de platino en un ácido fuerte y haciendo burbujear hidrógeno hasta la presión de 1 atmósfera*”. Dicha definición fue adoptada con algunas correcciones por la IUPAC en 1953 [14].

Electroquímica electródica. [3,4,5,6]

El siglo XX supuso la consolidación de la Electroquímica Iónica y el comienzo de la Electroquímica Electródica y la Cinética Electroquímica. Lo mismo que el estudio de la velocidad de reacción se conoce como cinética, en el caso de la electroquímica, estos estudios se conocen como cinética electroquímica o cinética de los procesos de electrodo.

La idea de que existe un potencial de contacto entre el electrodo y la disolución llevó al estudio de esa zona. En 1878 Helmholtz ya supuso que se trataba de una doble capa rígida con cargas de signo contrario similar a lo que ocurre en un condensador. Posteriormente con la influencia de las ideas de Ludwig Eduard Boltzmann (1844-1906) y más tarde en 1910 las de Louis George Gouy (1854-1926) y en 1913 las de David Leonard Chapman (1869-1958) se cambia ese modelo por otro que supone una capa rígida en electrodo

y una difusa en la disolución. En 1924 Otto Stern (1888-1969) describe un modelo que supone una capa rígida del tamaño de un ión adsorbida en el electrodo, teniendo en cuenta las ideas de Irving Langmuir (1881-1957, Premio Nobel de Química 1932), contemporáneo suyo, y a esa le seguiría otra difusa en la disolución.

Max Julius Louis Leblanc (1865-1943) en 1891 y William Augustus Caspari (1877-1951) en 1899 se dan cuenta de que para que se deposite un ión de un metal se necesita una tensión superior a la que se produce en el proceso inverso, es decir, cuando se rompe la estructura del metal y a ese exceso lo conocen como sobretensión. En 1905 Julius Tafel (1862-1918) descubrió que la sobretensión es función de la densidad de corriente y llegó a una ecuación que relaciona las dos magnitudes pero Tafel enfermó de pulmón, abandonó la ciencia y su trabajo cayó en olvido.

La estructura de la fase electrificada se pudo estudiar gracias a que en 1873 Gabriel Jonas Lippmann (1845-1921, Premio Nobel de Física 1908) descubrió la electrocapilaridad. Las mediciones mejoraron con Bohumil Kucera (1874-1921) que descubrió el electrodo de goteo de mercurio. En 1903 Kucera, en lugar de medir el movimiento del mercurio en un tubo capilar, medía el goteo del mercurio, que también es función de la tensión superficial, lo que se conoce como estudio dinámico, mientras que el de Lippmann era estático, y encontró que en las curvas de la tensión superficial en función del voltaje aparecen dos máximos, no acertando a explicar este hecho. Kucera formó parte del tribunal de la tesis doctoral de Jaroslav Heyrovsky (1890-1967, Premio Nobel de Química, 1959), que versaba sobre "Electroafinidad del aluminio", y, en el debate posterior, comprobó los conocimientos que éste tenía de electroquímica por lo que le animó a que buscara una explicación a esas curvas.

Heyrovsky comprobó que la adición a la disolución de cationes fácilmente reducibles producía fenómenos de inflexión en los voltajes próximos al voltaje de descomposición. Por otra parte, en lugar de medir el peso de las gotas de mercurio, decidió medir la intensidad de la corriente entre el mercurio que gotea y el fondo de la vasija. Las curvas obtenidas eran características de las disoluciones sobre las que se hace la electrolisis cualitativa y cuantitativamente. La posición de la curva en la gráfica, que viene dada por el voltaje, indica la composición cualitativa. La altura de la curva, que es la intensidad de corriente, determina la composición cuantitativa.

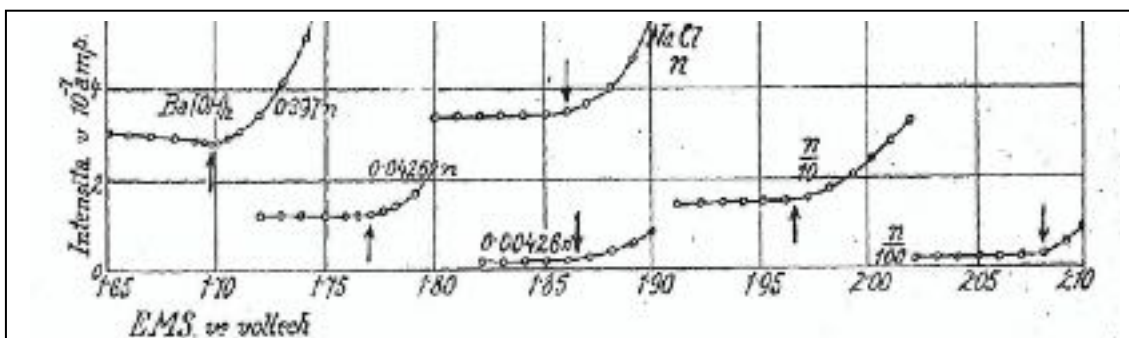


Fig.6. <http://knowledge.electrochem.org/estir/hist/hist-22-Heyrovsky-1.pdf>, p.259.

De esta forma nace la polarografía en 1925. Le da este nombre porque las curvas las registra sobre un cilindro rotatorio recubierto con papel fotográfico al que llama polarógrafo. En 1935 este método mejora con la utilización del osciloscopio. A Heyrovsky le impresiona sobremanera el uso del oscilógrafo como herramienta para representar las curvas. A partir de entonces, se convierte la polarografía en un importante método de análisis químico cualitativo y cuantitativo. Heyrovsky define la polarografía como *“la ciencia que estudia los procesos de electrodo que ocurren en torno al electrodo de gotas de mercurio”*. *“No incluye únicamente el estudio de las corrientes de sobrevoltaje, sino otras relaciones como las curvas intensidad de corriente-tiempo, por cada gota, curvas potencial-tiempo, fenómenos de electrocapilaridad y los flujos de corrientes en electrolitos y, como herramientas además del polarógrafo, se usan el microscopio, el galvanómetro y el oscilógrafo de rayos catódicos”* [15,16].

Mecánica Cuántica.

Las ideas de Tafel las revisan Tibor Erdey-Cruz (192-1976) y Max Volmer (1885–1965) en 1930 y John Alfred Valentin Butler (1899-1977) entre 1924-32 llegando a una ecuación que en condiciones concretas se reduce a la de Nernst. Pero mucho más importantes son los trabajos de Ronald Wilfred Gurney (1898- 1953) de 1931 y 1936 en los que aplica las ideas de la recién nacida Mecánica Cuántica a estos procesos. La segunda Guerra Mundial hizo que Gurney asustado por los problemas políticos en los que se había implicado, se fuera a Estados Unidos y muriera muy joven. Por otra parte, Butler se paso a las investigaciones en Biología y abandonó la Electroquímica.

Gurney [17] afirmó que la termodinámica había indicado en qué consiste el proceso de la electrólisis pero no podía explicar los mecanismos pero la Mecánica cuántica permitía abordar ese problema. Afirmaba *“Según la teoría de Sommerfeld los electrones de valencia de los metales son electrones libres por lo que podemos considerar los átomos de los cristales metálicos como iones. Si esta idea se aplica a un ánodo de cobre de un voltámetro, por ejemplo, podemos decir que cuando pasa corriente los iones de la red cristalina van dejando la superficie del metal y deslizándose dentro de la disolución. Esta idea es cierta para electrodos gaseosos y reversibles. En contraste con esto están los procesos de electrodo de una cuba en la que un ácido se descompone. En este caso los iones del electrolito se neutralizan por la pérdida captura de electrones produciendo oxígeno o hidrógeno neutros. Por lo que el fenómeno de la electrólisis tiene dos partes que se pueden abordar desde la Mecánica Cuántica. En la electrólisis nos encontramos con el complicado fenómeno del SOBREPOTENCIAL que proporciona una interesante prueba de nuestra teoría; por eso en este trabajo solo nos referiremos a los procesos irreversibles”*.

Continúa explicando cómo puede ocurrir la transferencia de electrones haciendo un estudio de niveles de energía.

A continuación existe un período en la electroquímica que solo merece atención la escuela rusa con Alexander Naumovich Frumkin (1895-1976),

fundador de la revista *Elektrokhimiya*, que siguió trabajando en procesos de electrodo introduciendo nuevas técnicas y nuevos conceptos pero que son poco conocidos, por una parte, por el aislamiento de Rusia y por otra, por el idioma.

En 1947, terminada la guerra mundial, la IUPAC organiza un simposio sobre Procesos de Electrodo donde se dan a conocer los representantes de la escuela Rusa y donde se afirma que Gurney estaba equivocado y Nernst tenía razón.

David Caldwell Grahame (1912-1958) entre 1941-1958 hizo estudios experimentales sobre la estructura de la doble capa que suponía la modificación de las ideas de Stern y que parecían prometedoras pero murió muy joven de un ataque al corazón.

Rudolph Arthur Marcus (1923) partiendo de las ideas de Gurney y con los conocimientos del momento, desarrollo el mecanismo del “efecto-túnel” y discutió su relación con el principio de Franck-Condon (regla de espectroscopia y química cuántica sobre las transiciones). Determinó cual es el complejo activado más estable a partir del máximo de la constante de la velocidad en función de la distancia. Mediante ese valor máximo calculó la mejor distancia de aproximación de los iones interactuantes que da el valor de la probabilidad de penetración del electrón coincidente con la energía más pequeña de activación en 1992. Entre 1956 y 1965 describió un mecanismo para la transferencia de electrones en el que consideraba que hay una superposición de los orbitales electrónicos de las moléculas reaccionantes en el complejo activado. Insistió en la importancia del principio de Franck-Condon (James Franck 1882-1964, Premio Nobel de Física 1925) y utilizó las energías de Max Born (1882-1970, Premio Nobel de Física 1954). Con la Mecánica estadística llegó a una teoría unificada de transferencia de electrón en reacciones homogéneas y en reacciones electroquímicas. Por todos sus trabajos se le concedió el Premio Nobel de Química en 1992 [18].

En los años sesenta apareció otra escuela importante con el alemán Heinz Gerischer (1919–1994) que fue poco conocida porque el idioma alemán había perdido fuerza y la traducción que se hizo de sus trabajos no estaba bien hecha y no se entendió. Con él vuelve la idea de potencial absoluto de electrodo. Fue el director de la tesis de Gerhard Ertl (1936) Premio Nobel de Química 2007 por ser el padre de la moderna química de superficie, que ha ayudado a explicar cómo las células de combustible producen energía sin contaminación, cómo los convertidores catalíticos limpian los escapes de los automóviles e incluso por qué se oxida el hierro.

Conclusiones.

De lo expuesto en este breve resumen queda patente la importancia del tema en el desarrollo de la ciencia química, de la industria y de la técnica, así como todas las posibilidades para desarrollarlo con los estudiantes mediante una enseñanza creativa que abarque distintas metodologías en las que se podrá

acudir a trabajos de tipo experimental para que sirvan para entender los teóricos.

Abarca diversas ciencias y un importante número de premios Nobel se han concedido por sus trabajos en este campo.

La Electrochemistry Society (ECS) tiene una página en la que recoge muchas de las publicaciones que a lo largo de la historia han tenido relación con este tema [19].

Referencias:

- [1] <https://electrochemistry.chemistryconference.org/market-analysis-pdfs/electro%20market160913072828.pdf>
- [2] J. J. Berzelius, Tratado de química mineral, vegetal y animal, traducido por R. Sáez Palacios y C. Ferrari Scardini, Imprenta de D. José María Alonso, Madrid, 1849.
- [3] M. T. Martín Sánchez, Evolución histórica de los procesos de electrodo y su repercusión en la metodología de la enseñanza (Tesis Doctoral, UNED, Madrid, 1990), (sin publicar).
- [4] M^aT. Martín Sánchez, M^aL. Fernández Castañón, A. Bernalte Historia de la Electroquímica, Química e industria: Q e I, ISSN 0033-6521, 9, 1991. 703-706.
- [5] M^aT. Martín Sánchez, M. Martín Sánchez, Doscientos años de la pila de Volta; An. Quím., Abril- julio 2001, 42-47.
- [6] J.R. Partington A History of Chemistry, Macmillan, New York, 1964 (Volume 4)
- [7] Galvani De viribus electricitatis in motu musculari, en Bononiensi Scentiarum et Artium Insittuto atque Academia Commetarii, VII, 363-418 , Modena 1792.
- [8] Volta Letter to Sir Joseph Bank Sur l'électricité excitée par le simple contact entre deux substances conductrices différentes , Phil.trans,1800,XC,403-431.
- [9] Volta On the Electricity excited by the mere Contact of conducting Substances of different kind, Philosophical Magazine 1800,VII,289-311.
- [10] W: R. Grove, "On Voltaic Series and the Combination of Gases by Platinum", Philosophical Magazine and Journal of Science , 1842,XXI ,417-420
- [11] J. J. Berzelius, W. Hisinger, Versuche über die Wirkung der electische Säule auf Salze und auf einige von ihren Basen:Gehlen"s N. Algern. J. Chem ,Leipzig,1803,I,115.
- [12] J.J Berzelius, W. Hisinger Ann Phys, 1807 , XXVII, 270-304.
- [13] M. Martín, G. Pinto, M^aT Martín: La nomenclatura en electroquímica: Aportaciones etimológicas y didácticas sobre los términos ánodo y cátodo, Revista española de física, ISSN 0213-862X, 2017 31 (4), 39-43.
- [14] J.H. Hildebrand Some Applications of the Hydrogen Electrode in Analysis, Research and Teaching. J. Am. Chem. Soc., 1913, 35 (7), 847-871.
- [15] M^a T. Martín Sánchez, M. Martín Sánchez, Corazón oscilante de mercurio An. Quím. 2010, 106(4), 304-310 18.
- [16] J. Heyrovský, Electrolysis with a dropping mercury cathode, The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science, 1923 Ser. 6, 45, 303-315.
- [17] R. W. Gurney Quantum Mechanics of Electrolysis and Quantum Mechanics of Electrochemistry, Proc. Royal Society, 134 A, 137-154 y 136 A ,378-391
- [18] R.A Marcus, R.A On the Theory of Electron-Transfer Reactions. VI. Unified Treatment for Homogeneous and Electrode Reactions. J. Chem. Phys. 1965, 43: 679-701
- [19] <http://knowledge.electrochem.org/estir/>

Manuela Martín Sánchez

Doctora en Ciencias Químicas.

Profesora jubilada Colaboradora Honorífica del Departamento de Didáctica de Ciencias Experimentales de la UCM.

Capítulo 8

Vacunas: historia de un gran paso para la sanidad.

A lo largo de la historia el hombre ha intentado encontrar protección contra distintas enfermedades. En el campo de la medicina, la vacunación ha supuesto un antes y después en el modo y calidad de vida de las personas.

Las primeras conjeturas de vacunación se remontan al siglo VII, cuando indios practicantes del budismo inhalaban o ingerían veneno de serpiente con el fin de ser inmunes a su picadura. Sin embargo, es en los textos chinos del siglo X “*El espejo dorado de la medicina*” y “*El tratamiento correcto de la viruela*”, dónde podemos ver los primeros indicios de la vacunación. En estos textos encontramos la explicación de distintas técnicas en la inoculación de la viruela. La inoculación, es introducir en una persona por medios artificiales los virus o bacterias que son los responsables de provocar la enfermedad [1].



*Fig.1. La inoculación, pulverizar costras desecadas e introducirlas mediante un tubo de bambú por los orificios nasales, ya se practicaba en China en el siglo XI.
(Imagen obtenida de delpensante.com)*

A pesar que está inoculación se utilizaba en el Imperio Turco y Asia, en Europa no hubo conocimiento de la variolización (inoculación de la viruela) hasta 1721, casi 700 años después, cuando la aristócrata inglesa Lady Mary Wortley Montagu acompañó a su marido embajador al Imperio Turco. Lady Mary, que había sufrido la pérdida de su hermano por la viruela, mostró un interés especial en la prevención de esta enfermedad y por ello, a la vuelta de

su viaje se trajo a Inglaterra la práctica de la variolización. A pesar de los prejuicios de la época, ella misma hizo inocular a sus hijos, demostrando ser así una revolucionaria para su época [2].

En 1757, el médico escocés Francis Home demuestra que el sarampión, que por aquella época era una de las enfermedades más comunes en niños de 15 años, estaba causada por un agente infeccioso que se hallaba en la sangre de los enfermos. Años más tarde en 1796, fue el inglés Edward Jenner, quien marcó un antes y un después en la historia de la vacunación, ganándose el título de “padre de la vacunación”.

¿Sabías qué? El origen de la palabra VACUNA procede de “vaca” porque las ordeñadoras que estaban en contacto con vacas enfermas de viruela se inmunizaban contra la viruela.



Fig.2. Retrato de Lady Mary Wortley Montagu (1689-1762).

Pongámonos en contexto, a finales del siglo XVIII la viruela, enfermedad provocada por el virus *Variola*, fue una plaga que no tenía ningún tratamiento o cura y que causó muchísimas muertes en Europa y América. La única manera de prevenir la viruela es la variolización, que consistía en inyectar a un sujeto sano algo de materia infectada procedente de un paciente con viruela, con la esperanza de que el sujeto pasara una enfermedad leve y así quedará protegido frente a ella. Sin embargo, esto no ocurría en todos los casos y mucha gente seguía falleciendo por la viruela. Sabiendo esto, Jenner observó que las ordeñadoras adquirían pústulas (heridas con pus) cuando estaban en contacto con las vacas, que por aquel entonces tenían la viruela de las vacas o *variola vaccina* en latín. Estas personas que sufrían pústulas eran inmunes a la viruela, es decir, no enfermaban. Con esta premisa, Jenner decidió probar a inocular la viruela de las vacas en una persona sana. El 14 de mayo de 1796 Jenner llevo a cabo dicho experimento, que se encuentra descrito en su publicación de 1798 “*Investigación sobre las causas y los efectos de la viruela vacuna*”:

“Para observar mejor cómo evolucionaba la infección, inoculé la viruela vacuna a un niño sano de ocho años. La vacuna procedía de una pústula del brazo de una ordeñadora, a quien había contagiado la vaca de su señor....

... el 1 de julio le inyecté materia virulosa que había extraído con anterioridad de una pústula humana. Se la apliqué profusamente mediante varios cortes y punturas, pero no dio lugar a ningún ataque de viruela.

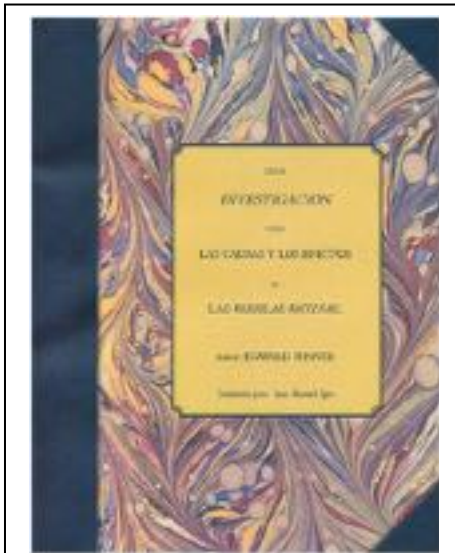


Fig.3. Imagen de la portada el libro "Una Investigación sobre las Causas y los Efectos de las Variolae Vaccinae" de Jenner, E. (Imagen obtenida de Axon)

En los brazos aparecieron los mismos síntomas que provocan las sustancias virulosas en los niños que han sufrido variola o viruela vacuna. Al cabo de unos meses, le volví a inocular materia virulosa, que en esta ocasión no produjo ningún efecto visible en el cuerpo".

E. Jenner

Después de intentar dar a conocer su éxito, Jenner sufrió constantes rechazos por parte de la Royal Society y de la Asociación Médica de Londres. Aunque todas las críticas tachaban el tratamiento de anticristiano, la vacunación acabó imponiéndose, y Jenner se convirtió en una celebridad hasta que en 1815 se retiró de la actividad científica. A pesar de los esfuerzos de Jenner, y de que el desarrollo de la vacuna de la viruela evitara

numerosas muertes, la viruela estuvo presente en África y Asia mucho tiempo después. Fue en 1966 cuando comienza el Programa de Erradicación de la Viruela por la Organización Mundial de la Salud (OMS). Y es en 1980, 24 años después del programa de vacunación se declaró oficialmente erradicada, convirtiéndose así en la primera enfermedad oficialmente eliminada de nuestro planeta [3,4].

La historia de la vacunación no termina aquí, el descubrimiento de Jenner fue tan revolucionario que el rey Carlos IV de España mandó al médico real Francisco Xavier de Balmis a las colonias españolas en el Nuevo Mundo, en la llamada "Real Expedición Filantrópica de la Vacuna". El 30 de noviembre de 1803 el navío María Pita salió del puerto de La Coruña con 22 niños desamparados, con el fin de ir vacunándolos a lo largo del viaje, de tal manera que hubiera pústulas en todo momento. Además de los niños, cabe destacar la presencia en el barco de Isabel Zendal, reconocida por la OMS como la **primera enfermera de la historia**.

Cuando la expedición llegó a Caracas, sólo uno de esos niños tenía pústulas, pero aun así Balmis inició la vacunación en Sudamérica, en la que es considerada la "primera expedición sanitaria internacional de

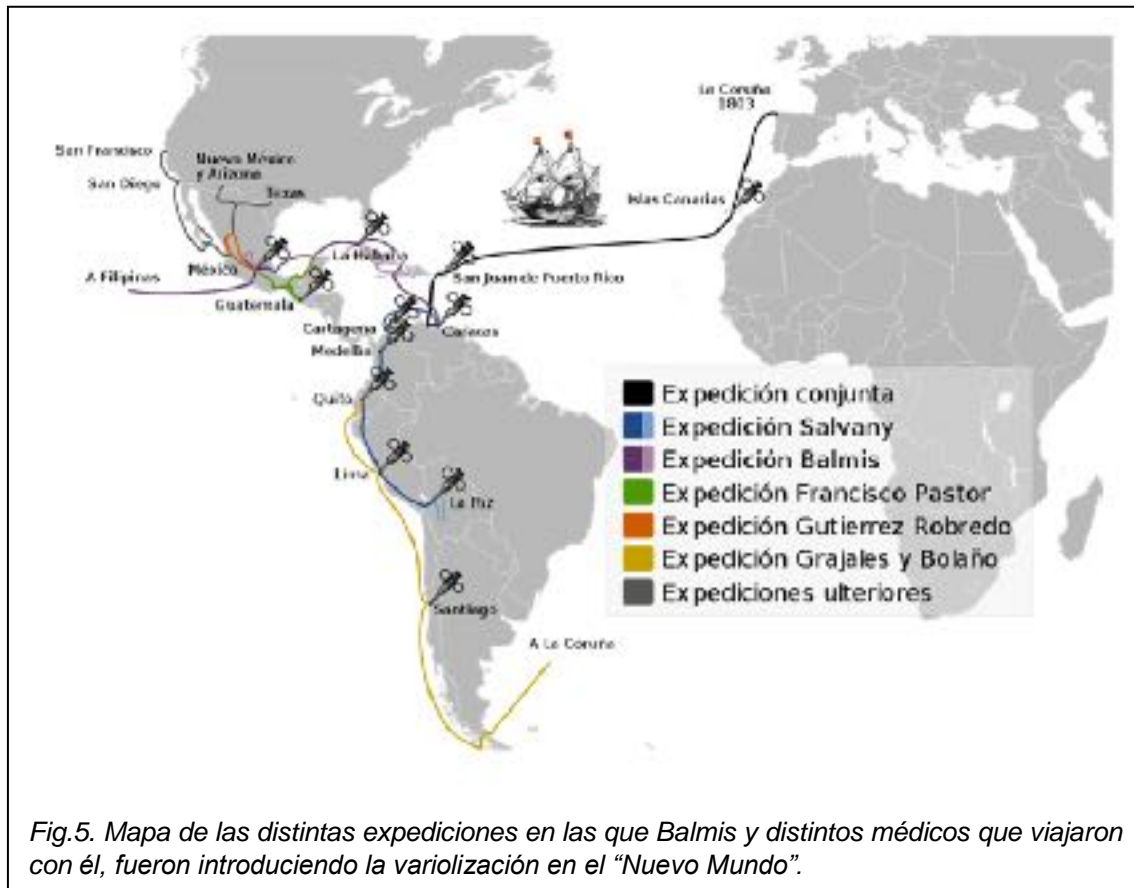


Fig.4. Fotografía de Janet Parker conocida como la última persona registrada en morir contagiada de viruela el 11 de septiembre de 1978, por fallos en la seguridad de un laboratorio de la Escuela de Medicina de la Universidad de Birmingham. (Imagen obtenida de Noticiaaldia)

la historia”. Tan importante fue la ayuda de Balmis, que conociendo que la vacuna no había alcanzado China, éste se fue a vacunar a la población de varias ciudades hasta llegar a la provincia de Cantón [5].

“No puedo imaginar que en los anales de la Historia se proporcione un ejemplo de filantropía más noble y más amplio que este”

E. Jenner



Obviamente este es un resumen, pero si contáis con la curiosidad de saber más detalles sobre la expedición de Balmis, la televisión española (TVE) realizó una serie llamada “22 Ángeles” en el que cuenta la expedición de manera más entretenida que los simples textos.

Si seguimos avanzando en la historia, a finales del siglo XIX se realizaron importantes logros en el campo de la microbiología y de la inmunología. Alguno de estos logros fueron llevados a cabo por Louis Pasteur, que además de ser el padre de “pasteurización” (aplicada hoy en día como norma de higiene en la industria alimentaria), fue el primero en producir una vacuna desarrollada en un laboratorio: la vacuna contra el cólera aviar. Pasteur debilitó por accidente las bacterias (un asistente se olvidó de cultivarlas en fresco) que inyectaba en pollos para estudiar la enfermedad del cólera aviar. Cuando los pollos sobrevivieron a la vacuna con las bacterias “débiles” o



Fig.6. Retrato de Luis Pasteur en su laboratorio.

atenuadas, cosa que no pasaba con las bacterias en fresco, Pasteur pensó que el factor que hizo que la bacteria fuera menos mortal había sido la exposición al oxígeno [1].

La tuberculosis era otra de las enfermedades que junto con la viruela y el cólera, causaba una gran mortalidad en la población. En microbiología, siempre nos hablan de Robert Koch y su “bacilo de Koch”

(1882) como uno de los principales pioneros en este campo. En 1905 Koch ganó el premio Nobel en fisiología/medicina por su investigación en la tuberculosis. Al final el “bacilo de Koch” sería conocido como *Mycobacterium tuberculosis*, responsable de la tuberculosis. Sólo tres años después de este descubrimiento, en 1885 ocurrieron dos grandes hechos históricos: la vacuna contra el cólera de Ferrán y el uso de la vacuna de la rabia en humanos. El primero de estos hechos se le otorga a Jaume Ferran i Clua, este médico desarrolló una vacuna contra el cólera aplicada con éxito en la epidemia de la región valenciana. Sin embargo, esta vacuna suscitó una fuerte oposición, hasta que finalmente su efectividad fue reconocida por la Academia de Ciencias de París, que le concedió en 1907 el Premio Bréant. El otro gran hito vuelve a estar en manos de Pasteur, vacunó 13 veces a un niño de nueve años que había sido mordido por un perro. Gracias a esto el niño nunca desarrolló la rabia, y su caso se consideró un éxito [6].



Fig.7. Retratos de Robert Koch, Jaume Ferran i Clua y Louis Pasteur, de izquierda a derecha.

Está claro que la velocidad de los hallazgos en el tratamiento de las enfermedades era cada vez mayor. En 1890 Shibasaburo Kitasato y Emil Von Behring demostraron que los sueros obtenidos de los conejos contenían una sustancia que prevenía la difteria, otra de las grandes plagas de la humanidad.

Posteriormente, demostraron que se podía curar la difteria si inyectaban el suero de un animal vacunado, a esto lo denominaron la “*terapia sérica*”. En 1901 Von Behring recibió el premio Nobel de medicina debido a su trabajo con la difteria. Como en muchos otros descubrimientos de la historia de la medicina, en la década de 1890 distintos investigadores analizaban al mismo tiempo los procesos para crear la protección contra la fiebre tifoidea. En Alemania, Richard Pfeiffer y Wilhelm Kolle demostraron que la inoculación de bacterias tifoideas muertas tenía como resultado la inmunidad humana contra la fiebre tifoidea. Ese mismo año, Almroth E. Wright demostró lo mismo en Inglaterra. Pero no fue hasta 1981, un siglo después, cuando se reconoció que Wright y Pfeiffer deberían haber compartido el crédito de la vacuna [7].

Durante los siguientes años ocurrieron distintos hitos históricos en la historia de la vacunación. Hay que destacar al médico del ejército de EE.UU, Jesse Lazear (dirigido por Walter Reed) que realizó experimentos que corroboraron la teoría expuesta por primera vez en 1881 por el cubano Carlos Finlay, la fiebre amarilla se transmitía por los mosquitos. Otro de los hitos importantes de la vacunación, fue el descubrimiento de la vacuna contra la tuberculosis. Para evitar confusiones, Koch



Fig.8. Pintura de Robert Thom que muestra miembros de la comisión de fiebre amarilla en Cuba.

descubrió el bacilo, como se menciona anteriormente, pero la vacuna no fue descubierta hasta 1921 por Albert Calmette y Camile Guerin. Hoy en día, además de ser usada como medio de prevención de la tuberculosis, la vacuna BacillusCalmetteGuerin (BCG) se emplea como tratamiento en la prevención de reaparición tumoral en el cáncer de vejiga. En Francia, Gaston Ramon un veterinario del Instituto Pasteur, desarrolló el toxoide de la difteria. De manera independiente, en los Wellcome Research Laboratories de Londres, el doctor Alexander Thomas Glenny hizo lo mismo en 1923 [1,7].



Fig. 9. Escultura de Balto en Central Park (Nueva York). (Historia de Nueva York).

Seguro que si se menciona el nombre de Balto, se viene a la cabeza el perro de la famosa película de dibujos animados “Balto: la leyenda del perro esquimal”, basada en un hecho histórico que no se puede dejar pasar. En 1925 Alaska sufrió un brote de difteria, Curtis Welch, un médico de Nome (Alaska), envió un telegrama solicitando la entrega de antitoxinas por correo. El envío de estas toxinas se tendría que

realizar en trineo tirado por perros para llegar a Nome, y el viaje no fue fácil, los perros tuvieron que luchar contra las bajas temperaturas, hubo que hacer paradas técnicas para mantener caliente el suero. Por su gran esfuerzo, Togo y Balto, dos de los perros de trineo, recibieron atención nacional. Una estatua de Balto sigue en pie en Central Park en la ciudad de Nueva York [7].

A partir de aquí la lista de vacunas se amplía hasta conseguir lo que entendemos hoy en día como “calendario de vacunación”. Destacando por su importancia la vacuna contra la fiebre amarilla (Theiler, 1936), tos ferina (Kendrik y Elderding, 1939), poliomielitis (Salk, 1955), sarampión (Enders, 1960), rubéola (Séller, 1962), paperas (Hilleman, 1965), gripe (Hilleman, 1968), varicela (Takahashi, 1974), meningitis (Hilleman, 1974), hepatitis B (Hilleman, 1979), hepatitis A (Hilleman, 1995) y el virus de papiloma humano (Merck, 2006) [1,7].



Fig.10. Fotografía de Maurice Hilleman, conocido como el hombre de las 40 vacunas.

Desde el inicio de la vacunación, han existido grupos o movimientos en contra de esta práctica. Uno de estos movimientos fue la creación de la Sociedad Anti-Vacunación de América en 1879, y la Liga Anti-Vacunación de Nueva York en 1885. En noviembre de 1904, en respuesta a una campaña pública de vacunación, ciudadanos y cadetes militares de Rio de Janeiro se levantaron en la Revolta da Vacina. Los disturbios empezaron el día en que la ley de vacunación entró en vigor. Más recientemente, algunas actuaciones de movimientos anti-vacunas han llevado a un rebrote en el número de casos de determinadas enfermedades infecciosas. Un ejemplo lo tenemos en Europa, dónde el sarampión está cada vez más extendido como consecuencia de la falta de vacunación. Otro ejemplo es lo sucedido en Nigeria, donde los líderes religiosos recomendaron la no-vacunación de la poliomielitis, hecho que ha tenido como consecuencia la reaparición de esta enfermedad en países libres de poliomielitis [7,8].

En el informe de la OMS de 2019, dónde se enumeran los diez problemas de atención en salud, destaca la renuencia o reticencia a vacunarse. La vacunación previene de 2 a 3 millones de muertes por año y podrían evitarse otros 1.5 millones si se mejorara la cobertura mundial de vacunas [9].

El mundo se ha enfrentado y se enfrenta a múltiples retos de salud. Las vacunas son una de las herramientas esenciales de la salud pública, con grandes beneficios para la salud y el bienestar de la población. No desperdiciemos lo que muchos han tardado siglos en conseguir.

Referencias:

- [1] Berdasquera Corcho, D; Cruz Martínez, G; Suárez Larreinaga, C.L. 2007. (2000). La vacunación: Antecedentes históricos en el mundo. *Revista Cubana de Medicina General Integral*. 16. 375-378.
- [2] Dinc, G; Ulman YI. 2007. The introduction of variolation 'A La Turca' to the West by Lady Mary Montagu and Turkey's contribution to this. *Vaccine*, 25(21):4261-5.
- [3] Riedel, S.2005. Edward Jenner and the History of Smallpox and Vaccination. *Proceedings (Baylor University.Medical Center)*, 18(1), 21-5.
- [4] Bazin, H. 2000. The Eradication of Smallpox 1st Edition, Edward Jenner and The First and Only Eradication of a Human Infectious Disease. *Academic Press*. ISBN 978-0-12-083475-4.
- [5] Tuells, J. 2002. Francisco Xavier Balmis (1753–1819), a pioneer of international vaccination. *J Epidemiol Community Health*. 56(11): 802.
- [6] Metchnikoff, E.2006. The Founders of Modern Medicine: Pasteur, Koch, Lister. Delanco, N.J. *The Classics of Medicine Library Division of Gryphon Editions*.
- [7] The College of physicians of Philadelphia. *The History of vaccines* (<https://www.historyofvaccines.org/>). Consultado en Enero 2019.
- [8] López Santamaría, M.A. (2015). *Los movimientos antivacunación y su presencia en internet*. Ene, 9(3) <https://dx.doi.org/10.4321/S1988-348X2015000300011>
- [9] WHO. 2019. *Ten threats to global health in 2019* (<https://www.who.int/es/emergencias/ten-threats-to-global-health-in-2019>). Consultado en Enero 2019.

Carolina Rubio Martín

Graduada en Biología, especialidad Microbiología.

Investigador Pre-Doctoral en el Departamento de Oncología Molecular en el Centro de Investigaciones Energéticas, Medioambientales y Tecnológicas (CIEMAT).

Capítulo 9

El Electromagnetismo: una historia de grandes pasos.

En el presente capítulo se hará un breve repaso a la Historia del Electromagnetismo como ciencia, desde sus orígenes hasta la aparición de artilugios eléctricos como bombillas, motores, tendidos eléctricos, etc.

En primer lugar, antes de comenzar dicho repaso, es importante señalar que el conocimiento de los principios que rigen el electromagnetismo ha permitido a los seres humanos desarrollar toda una serie de inventos con aplicaciones prácticas (algunos de ellos ya mencionados), que han sido fundamentales en la profunda transformación que ha experimentado el modo de vida de los humanos durante los últimos dos siglos aproximadamente, especialmente en los países desarrollados. A dicho proceso de transformación disruptiva se le puede llamar, sin temor a sobreestimar su impacto e influencia, la Revolución Eléctrica.

Por lo tanto, el conocimiento de la Historia del Electromagnetismo nos acerca de manera decisiva a nuestra propia Historia, y nos permite entender mejor el mundo en el que vivimos actualmente, y la manera en la que hemos llegado hasta aquí.

Hecha esta pequeña -pero notoria- aclaración, comenzamos ahora sí con el repaso a la Historia del Electromagnetismo. Tenga el lector presente que esta exposición estará necesariamente simplificada, y que muchas contribuciones al desarrollo de esta ciencia no serán tratadas en la misma manera que merecerían, dado que el objetivo es servir como introducción al tema y discutir los principales hitos en la construcción de una de las grandes teorías clásicas de la Física (en otras palabras: la Historia completa del Electromagnetismo no cabe en unas pocas páginas).

Los primeros descubrimientos.

Las observaciones de fenómenos eléctricos, tales como tormentas eléctricas o peces eléctricos (como por ejemplo las anguilas), deben de haberse producido desde tiempos inmemoriales; muy probablemente, a tales fenómenos se les atribuiría un carácter mágico o mitológico. Los primeros indicios acerca de la existencia de fuerzas eléctricas y magnéticas se remontan sin embargo a la Antigüedad, y suelen atribuirse (¡oh, sorpresa!) a los griegos. En torno al año 600 a.C., el filósofo griego Tales de Mileto, describió cómo el ámbar (resina de árboles fosilizada) atraía pequeños y ligeros objetos cuando era frotado con

pieles de animales. La palabra que empleaban los griegos para designar el ámbar era “*elektron*”, y de ahí procede el nombre de “electricidad”. Los griegos también sabían que la magnetita, un mineral compuesto por óxidos de hierro, atraía trozos de hierro. La palabra griega para designar la magnetita es “*magnetis lithos*”, y de ahí procede el nombre de “magnetismo”.

Por otra parte, el uso de la brújula magnética es también antiquísimo (recordemos que una brújula es una aguja de hierro –o de magnetita- imantada que se alinea según el campo magnético de la Tierra, señalando de esta forma sus extremos el polo norte y el polo sur terrestres, aproximadamente). Se suele atribuir a los chinos la invención de la brújula magnética, posiblemente durante el siglo III a.C. Es probable que los vikingos también la inventaran de manera independiente, si bien más tardíamente (alrededor del siglo I d.C.). Por supuesto, ambas civilizaciones se habrían beneficiado inmensamente de tal invención en lo concerniente a la navegación, aunque parece que ésta no fue la aplicación original en la antigua China. En Europa, en contraste, la primera referencia conocida a la brújula data de principios del siglo XIII d.C.

Los orígenes como Ciencia.

En 1269, el francés Petrus Peregrinus de Maricourt colocó un fino rectángulo de hierro sobre una esfera de magnetita en diferentes posiciones, y observó que se alineaba siempre en meridianos análogos a los de la Tierra, lo cual le llevó a definir el concepto de polos magnéticos. Descubrió también que cuando un imán es cortado en dos, cada una de las piezas resultantes también tiene dos polos. Y observó que polos semejantes se repelen y que polos opuestos se atraen. Recogió el resultado de sus investigaciones en su “*Epistola de Magnete*”, que hoy en día constituye el primer (pequeño) tratado conocido sobre el tema.

Hasta finales del siglo XVI sin embargo, no se produjeron apenas nuevos avances. En el año 1600, William Gilbert, médico de la reina Isabel I de Inglaterra, que había dedicado años a estudiar de manera sistemática el magnetismo (y en menor medida, la electricidad estática), publicó “*De Magnete*”; suele considerarse la publicación de esta obra como el origen del Electromagnetismo como disciplina científica. Entre otras cosas, Gilbert concluyó que la Tierra debía ser un inmenso imán. Esta notable afirmación contrasta con la visión aceptada en la época, según la cual el magnetismo se debía a la atracción de los polos de la esfera celeste.

El siglo XVIII: la Electroestática.

El siglo XVII vio nacer los generadores electrostáticos, aparatos capaces de almacenar carga eléctrica, generalmente por fricción. El primer generador electrostático fue el construido por Otto von Guericke en 1663, aunque numerosos diseños distintos aparecieron con posterioridad. Entre todos ellos, destacaron significativamente las botellas o jarras de Leyden, inventadas de manera independiente por Peter von Musschenbroek (1746), un científico

holandés de la ciudad de Leiden, y el alemán Ewald Georg von Kleist (1745). La botella de Leyden era una botella de vidrio sellada, llena de agua hasta la mitad, que contenía un cilindro conductor que sobresalía por la abertura de la botella. La carga se le comunicaba al cilindro mediante un generador de fricción. William Watson, médico y científico inglés, añadió un recubrimiento metálico tanto en el interior (en contacto con el cilindro conductor) como en el exterior, lo que aumentó enormemente la capacidad de las botellas y constituyó la invención del primer condensador eléctrico; hoy en día los condensadores se emplean de manera rutinaria en infinidad de aparatos electrónicos.

Los generadores electrostáticos podían acumular grandes cantidades de carga, y jugaron un papel fundamental en el nacimiento y desarrollo, a principios del siglo XVIII, del estudio científico de la electricidad, ya que posibilitaron la investigación de las fuerzas eléctricas entre objetos cargados, en situaciones en las que las cargas son estáticas. Así, en 1785 Charles Coulomb determinó experimentalmente la ley que lleva su nombre, según la cual la fuerza eléctrica entre dos cargas (o entre los polos de dos imanes) disminuye con el cuadrado de la distancia, y es proporcional al valor de cada carga. Para ello inventó la balanza de torsión, un importante instrumento que sirve para medir fuerzas que son débiles. La ley de Coulomb había sido sin embargo ya demostrada por el científico escocés John Robinson en 1769 (aunque de manera menos concluyente), y al parecer también por el extravagante científico inglés Henry Cavendish (quien sin embargo no publicó, fiel a su costumbre, sus resultados). Por su parte, los matemáticos Siméon-Denis Poisson (en 1813), George Green (en 1828) y Carl Friedrich Gauss (en 1835) realizaron importantísimas contribuciones matemáticas a la teoría del Electromagnetismo, y de esta forma, a principios del siglo XIX quedaron establecidos los principios de la electrostática.

La pila de Volta.

En el año 1800 se produjo un descubrimiento crucial, que cambiaría el rumbo de la Historia. Ese año, Alessandro Volta escribió una carta al presidente de la Royal Society de Londres, sir Joseph Banks. Junto a la carta envió un aparato compuesto por discos de cinc y de cobre apilados alternativamente uno sobre otro en pares (cinc, cobre, cinc, cobre...), y separados los pares entre sí por paños empapados en una solución salina (salmuera), también en forma de disco (Fig.1). Cuando se conectaban los extremos de esta pila de discos mediante cables conductores, se producía una corriente continua (esta corriente se manifestaba mediante el calentamiento de los materiales conductores, o la separación de compuestos químicos, como el agua, en sus elementos constituyentes)⁸.

⁸El hecho de que la electricidad podía ser transmitida por un material conductor se conocía gracias al descubrimiento llevado a cabo por Stephen Gray en 1729, que fue capaz por ejemplo de transmitir una corriente a lo largo de un tallo de cáñamo de 150 metros. Había buenos motivos por lo tanto para sospechar la existencia de una verdadera "corriente material" a través de los materiales conductores.



Fig. 1. Pila de Volta. Fuente: https://commons.wikimedia.org/wiki/File:Pila_di_Volta.jpg

Este artilugio, que ha pasado a la historia como “la pila de Volta”, revolucionó el estudio de la electricidad, ya que permitió disponer de una fuente de corriente continua duradera. Hasta entonces, se disponía solamente de generadores de carga tales como las botellas de Leyden, que si bien permitían acumular una gran cantidad de carga (lo cual era apropiado para el estudio de la electrostática), se descargaban muy rápidamente, y no permitían por lo tanto el estudio de las corrientes eléctricas.

La pila de Volta fue por lo tanto un invento de la mayor relevancia, sin el cual no se podrían haber producido los descubrimientos posteriores de Ampère, Faraday, etc. sobre el electromagnetismo. Tanto es así, que su valía fue reconocida de inmediato. Francesc Aragó, quien también realizó

contribuciones notables al electromagnetismo, escribió: “*La producción de electricidad por el mero contacto de metales diferentes es uno de los hechos más grandes y establecidos de la Ciencia Física... y ahora estos medios están en manos de todos los experimentadores, y es gracias al genio de Volta que lo ha hecho posible*”. O también: “[es] *el instrumento más maravilloso que el hombre nunca ha inventado, sin exceptuar el telescopio y la máquina de vapor*” [1]. Volta incluso llegó a presentar su invento en una sesión particular ante el propio Napoleón I Bonaparte, quien, impresionado, le otorgó la medalla de oro al mérito científico y le nombró conde y senador de Lombardía.

Sin embargo, la pila de Volta tenía un problema: el peso de los discos comprimía el paño húmedo, lo cual provocaba cortocircuitos. Para solucionar este inconveniente, William Cruickshank, químico de la Royal Military Academy de Woolwich, ideó una sencilla pero decisiva mejora, introducida a principios de 1802: colocó los componentes de la pila en horizontal en el interior de una caja de madera, y sustituyó la solución salina por ácido sulfúrico.

La mejora introducida por Cruickshank permitió construir baterías cada vez más grandes, con un mayor número de discos o placas metálicas (algo que no habría sido posible con el diseño original de Volta), las cuales proporcionaban voltajes e intensidades de corriente mayores. De hecho, a principios del siglo XIX comenzó una carrera tecnológica entre Francia e Inglaterra por ver quién construía la pila más grande [1]. En 1802, Humphry Davy había construido una gran pila que le había permitido aislar por primera vez mediante electrólisis diferentes elementos químicos, tales como el sodio, el potasio, ... (ver el capítulo de Electroquímica). Sin embargo, hacia 1808 la pila ya estaba muy desgastada, y se hizo necesario construir una nueva. Por aquel

entonces, Joseph Louis Gay-Lussac y Louis Jacques Thénard estaban inmersos en la construcción de una batería de grandes dimensiones en Francia. La Royal Institution de Londres, a la que pertenecía Humphry Davy, andaba escasa de fondos, e inició una campaña (apelando al nacionalismo y a la grandeza de la nación de Inglaterra) para recaudar fondos para construir una gran batería antes que los franceses. Finalmente, los ingleses consiguieron su objetivo en 1810, con la construcción de una inmensa batería que generaba 2.200 voltios, mientras que la pila de los franceses no llegaría hasta 1813.

La unificación de la electricidad y el magnetismo.

El descubrimiento de Oersted.

Como ya se ha señalado, la invención de la pila de Volta permitió el comienzo de los estudios sobre las corrientes eléctricas. Así, en 1820 Hans Christian Oersted, químico y físico danés, realizó un descubrimiento crucial, que desataría un torbellino de investigaciones acerca de la relación entre la electricidad y el magnetismo. Oersted creía firmemente en la interconexión entre ambos fenómenos desde hacía más de una década, en un momento en el que no existían indicios de tal interdependencia y pocos científicos podían siquiera sospecharla. Esta convicción filosófica le llevó a demostrar que una corriente eléctrica provoca la desviación de una aguja imantada. En otras palabras: las corrientes eléctricas se comportan como imanes, ejerciendo una fuerza sobre las agujas imantadas y desviándolas de su posición natural, orientada según el campo magnético terrestre. En concreto, Oersted demostró que una corriente eléctrica continua que circula por un conductor rectilíneo ejerce una fuerza sobre una aguja imantada que tiende a alinear ésta a lo largo de una circunferencia centrada en el conductor (la alineación es mayor cuanto mayor sea la intensidad de la corriente, debido a que hay que compensar la fuerza ejercida por el campo magnético de la Tierra). Era la primera demostración de que la electricidad y el magnetismo estaban conectados de algún modo, y supuso un verdadero hito en la Historia del Electromagnetismo.

Las aportaciones de Ampère (y Arago, Biot y Savart).

A raíz de este importante descubrimiento, los científicos de la época comenzaron un gran número de experimentos, con resultados notables en un período de tiempo muy breve. Entre todos ellos, destacó especialmente el francés André-Marie Ampère. Ampère formuló una teoría electromagnética según la cual todos los efectos magnéticos, incluidos los de los imanes, eran el resultado de electricidad en movimiento (esto es, de corrientes eléctricas—que serían internas en el caso de los imanes—), y no de un “fluido magnético”, tal y como se creía entonces. Asimismo, llevó a cabo notables descubrimientos en el plano experimental que confirmaron sus intuiciones. Descubrió que la propia pila voltaica, cuando sus extremos eran conectados, provocaba también la deflexión de una aguja imantada del mismo modo que lo hace un cable por el que circula corriente; para llevar a cabo este descubrimiento, inventó el galvanómetro, un sensible aparato que se sigue utilizando hoy en día para medir la corriente eléctrica (en base precisamente a la deflexión causada por la corriente en una aguja imantada); descubrió que dos cables conductores

rectilíneos por los que circula corriente se atraen o se repelen mutuamente según la corriente vaya en el mismo o en distinto sentido en ambos (¡y que esta atracción no es una atracción eléctrica!, un buen indicio de que las corrientes producen efectos magnéticos); en realidad, este último descubrimiento estuvo precedido por la observación de que un cable enrollado en forma de hélice por el que circula una corriente (es decir, un solenoide) se comporta exactamente como un imán, atrayendo y repeliendo otros imanes; esta espectacular revelación vino a su vez precedida además por una predicción teórica suya, según la cual la acción de los imanes se debía precisamente a corrientes circulares alrededor de su eje longitudinal, por lo que este descubrimiento no fue casual, sino buscado (para contrarrestar los efectos debidos a la extensión del cable en la dirección de dicho eje, Ampère dobló sobre sí mismos los extremos del cable y los extendió hacia el centro de la hélice); este descubrimiento supuso también de facto la invención de los electroimanes⁹, que hoy en día se utilizan de manera rutinaria en una gran variedad de dispositivos eléctricos, y que constituyeron la clave para la invención del telégrafo (se utilizan también por ejemplo en los aceleradores de partículas para dirigir a éstas); demostró además, en colaboración con François Arago, que un electroimán podía provocar la magnetización de una aguja de hierro colocada a lo largo de su eje; y que el campo magnético de la Tierra ejerce una fuerza sobre un circuito circular plano que tiende a alinearlo perpendicular a los meridianos magnéticos terrestres.

En definitiva, estableció de manera esclarecedora una serie de hechos experimentales y reveló la naturaleza eléctrica de los campos magnéticos. Elaboró además toda una teoría del Electromagnetismo que publicó en 1825 en su importante obra "*Mémoire sur la théorie mathématique des phénomènes électrodynamiques uniquement déduite de l'expérience*". En un principio pensó que la electricidad era un fluido¹⁰, pero ante los repetidos intentos en vano de encontrar evidencias en una serie de ingeniosos experimentos llevados a cabo por su amigo y físico Augustin Fresnel, asumió la idea de éste de que la

⁹En 1824, William Sturgeon introdujo los núcleos de hierro (cuya función es aumentar la intensidad del campo magnético) en los electroimanes, al enrollar un cable de cobre alrededor de un núcleo en forma de herradura. Esta mejora fue determinante en los experimentos posteriores que llevaron al descubrimiento de la inducción electromagnética por parte de Michael Faraday y Joseph Henry. El propio Henry construyó electroimanes de gran potencia, capaces de levantar casi una tonelada, gracias a que utilizó seda para aislar el cable conductor, lo que le permitió enrollar el cable de manera mucho más compacta y potenciar así aún más el campo magnético.

¹⁰La idea de que la electricidad era un fluido estaba bastante aceptada durante los siglos XVII, XVIII y XIX. Charles François de Cisternay Du-Fay defendió en 1733 que existían dos fluidos eléctricos, y que en general la materia era neutra, pues contenía igual cantidad de ambos; cuando por fricción, un cuerpo adquiría más de un fluido que de otro, entonces atraía o repelía otros objetos igualmente cargados. Benjamin Franklin, inventor del pararrayos tras su famoso experimento con la cometa, pensaba por el contrario que existía un único fluido eléctrico, y que un exceso o un defecto de este fluido en un cuerpo provocaba que los cuerpos estuvieran cargados positiva o negativamente. Las ideas de Franklin, introducidas por primera vez en 1747, dominaron el estudio de la electricidad durante 100 años aproximadamente. Hoy sabemos que ninguna de las dos opiniones es totalmente correcta. En primer lugar, la electricidad no es un fluido. Después, por un lado, Franklin tenía razón en que un único tipo de carga es el que se transmite en las corrientes eléctricas (los electrones). Por otro lado, Du-Fay tenía razón en que existen dos tipos de carga eléctrica, que hoy denominamos positiva y negativa.

corriente eléctrica se debía a una “corriente molecular” de numerosas y pequeñas partículas cargadas, que se movían dentro del conductor. Esta idea, que a la postre se demostraría cierta tras el descubrimiento del electrón por parte de J.J. Thomson en 1897, fue cuestionada por muchos científicos de la época, incluyendo a Michael Faraday [2].

Por su parte, François Arago demostró que un cable por el que circula corriente eléctrica es capaz de atraer limaduras de hierro como si fuera un imán y magnetizar de manera permanente una aguja de coser. Otra aportación importante fue la de Jean-Baptiste Biot y Félix Savart quienes, utilizando un cable conductor y pequeños cilindros de acero magnetizado suspendidos en hilos de seda de gusano, establecieron la ley que lleva su nombre, según la cual la acción de un elemento muy pequeño de cable por el que circula corriente, sobre un pequeño imán, se produce en la dirección perpendicular al cable y a la línea que une perpendicularmente el punto y el cable.

Todas estas notables aportaciones y descubrimientos tuvieron lugar durante el propio año 1820, a principios del cual Oersted había anunciado su hallazgo. Así, en apenas un año se pasó fugazmente de una situación en la que no existía ningún indicio sobre la conexión entre electricidad y magnetismo, a otra en la que buena parte de dicha conexión había sido revelada con acierto (al menos en lo referente a la magnetostática, esto es, al estudio de los campos magnéticos creados por corrientes que no varían en el tiempo). Faltaban aún sin embargo unos cuantos hallazgos sorprendentes por llegar.

El descubrimiento de la inducción electromagnética.

Hacia 1822, estaba claro que las corrientes eléctricas producían efectos magnéticos. Una pregunta natural entonces era si los imanes -incluidos los electroimanes- podían producir efectos eléctricos. Pese al interés y los múltiples intentos por demostrar tal conjetura, a comienzos de la década de 1830 ningún resultado en positivo había sido hallado. Entonces, en 1831, Michael Faraday enrolló dos bobinas de cobre en torno a un anillo de hierro, a modo de transformador (Fig.2), conectó una de ellas a una pila voltaica, y la otra a un galvanómetro muy sensible. Hizo pasar corriente por la primera de las bobinas y observó en ese momento que la aguja del galvanómetro oscilaba muy fugazmente cuando conectaba la pila, y que la aguja volvía a oscilar rápidamente, esta vez en sentido contrario, cuando desconectaba la pila. El efecto duraba apenas una fracción de segundo, y desaparecía en cuanto la corriente se estabilizaba. Posiblemente ése era el motivo por el

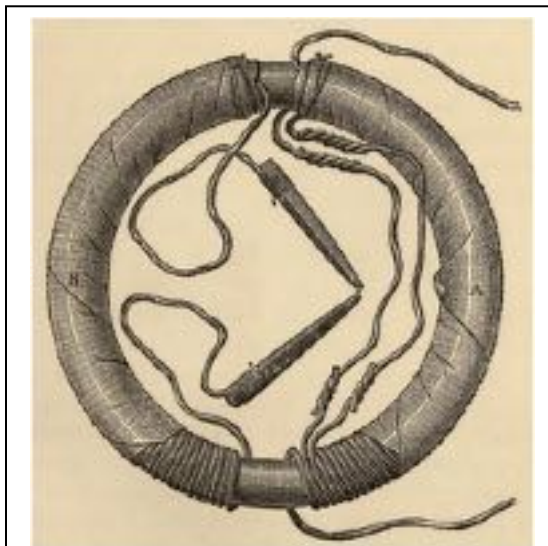


Fig.2. Grabado del transformador utilizado por Faraday para descubrir la inducción electromagnética. Fuente: https://commons.wikimedia.org/wiki/File:Faradays_transformer.png

cual no había sido descubierto con anterioridad: ¿no era la propia corriente eléctrica la que causaba la inducción, sino la variación de la corriente!

Faraday se dio cuenta de este hecho sutil, y llevó a cabo una sistemática serie de experimentos para dilucidar el fenómeno de la inducción. Descubrió así que, si acercaba o alejaba un imán a una bobina, o una bobina a un imán, o si acercaba o alejaba una bobina por la que circulara corriente a otra, también se inducía corriente eléctrica en la bobina. Descubrió también que la corriente inducida era tanto mayor cuanto más rápido fuera el movimiento, más potente el imán y mayor el área de la bobina. En definitiva, describió de manera exhaustiva el fenómeno de la inducción electromagnética.

Ésta fue sin duda la aportación a la Ciencia y al mundo más determinante de Faraday (entre las numerosísimas que hizo [3]): la inducción electromagnética permite la generación de la electricidad que consumimos, en centrales térmicas, hidroeléctricas, eólicas y nucleares, y ha posibilitado lo que podríamos llamar la Revolución Eléctrica (en analogía con la Revolución Industrial, acontecida a raíz de la invención de la máquina de vapor).

Es interesante, llegados a este punto, destacar que todo comenzó con unas modestas y fugaces deflexiones de la aguja de un galvanómetro. Este hecho ilustra perfectamente el funcionamiento de la Ciencia Básica, en contraste con la Ciencia Aplicada. Se dice que, en una ocasión, Faraday, al ser preguntado sobre la utilidad del descubrimiento que acababa de realizar, respondió: «¿Y para qué sirve un recién nacido?» [4].

También Joseph Henry, profesor universitario estadounidense, realizó en los mismos años que Faraday experimentos muy similares, que al parecer también lo condujeron al hallazgo de la inducción electromagnética, si bien Faraday publicó antes los resultados. Henry descubrió además la autoinducción (descubrimiento que sí se le reconoce), mejoró y popularizó los electroimanes (llegó a construir uno capaz de sostener masas cercanas a una tonelada) e inventó el telégrafo en 1831, aunque no lo patentó (posteriormente ayudaría a Samuel Morse en la construcción de su telégrafo, entre Baltimore y Washington). Las aportaciones de Henry lo convierten por lo tanto en una figura descomunal en la Historia de la Física. Una figura que, no obstante, se ha visto eclipsada por la de Michael Faraday, pero que es de justicia destacar.

La unificación de luz y electromagnetismo: la síntesis electromagnética.

Con el descubrimiento de la ley de inducción por parte de Faraday (con el permiso de Henry), y los descubrimientos realizados con anterioridad por Oersted, Ampère, Arago y compañía, se había logrado hallar una bonita interdependencia entre la electricidad y el magnetismo: los fenómenos eléctricos producían efectos magnéticos (más aún, eran la única fuente de los mismos); al mismo tiempo, estos fenómenos magnéticos producían efectos eléctricos.

La teoría del Electromagnetismo clásico iba tomando forma. Gran cantidad de hechos experimentales habían sido probados, y también

interpretados a la luz de consideraciones teóricas, y se habían incluso diseñado y fabricado artilugios como motores y generadores eléctricos. Existían sin embargo discrepancias en la interpretación sobre la manera en la cual un cuerpo cargado ejercía su ‘influencia’ electromagnética sobre otro. Por un lado, un gran número de científicos consideraban que se producía una “acción a distancia” entre las cargas, tal y como sugerían las ecuaciones, y como se pensaba que se comportaba también la gravedad entre cuerpos con masa. Por su parte, Faraday creía que las cargas distorsionaban de algún modo el espacio a su alrededor, creando una cierta “tensión”, y que esta distorsión o tensión era la que provocaba las fuerzas eléctrica y magnética, que se podían explicar y representar mediante las denominadas “líneas de fuerza”. Éstas constituirían una especie de red que llenaría el espacio alrededor de los cuerpos cargados, y servirían como representación gráfica para indicarla intensidad, la dirección y el sentido de las fuerzas eléctrica y magnética (Fig.3). Para Faraday, la existencia de tales líneas reflejaba la integridad e interconexión de todas las cosas, en tanto que obras de Dios (cabe recordar que Faraday pertenecía a una minoritaria y ferviente corriente religiosa de la Iglesia presbiteriana, los sandemanianos). Hoy en día se aceptan las ideas de Faraday, se entiende el electromagnetismo (y también otras teorías físicas) en términos de campos, y las líneas de fuerza se utilizan como representación de los campos eléctricos y magnéticos. Esto es así gracias a la denominada síntesis electromagnética llevada a cabo por James Clerk Maxwell, profesor escocés del King’s College por aquella época.

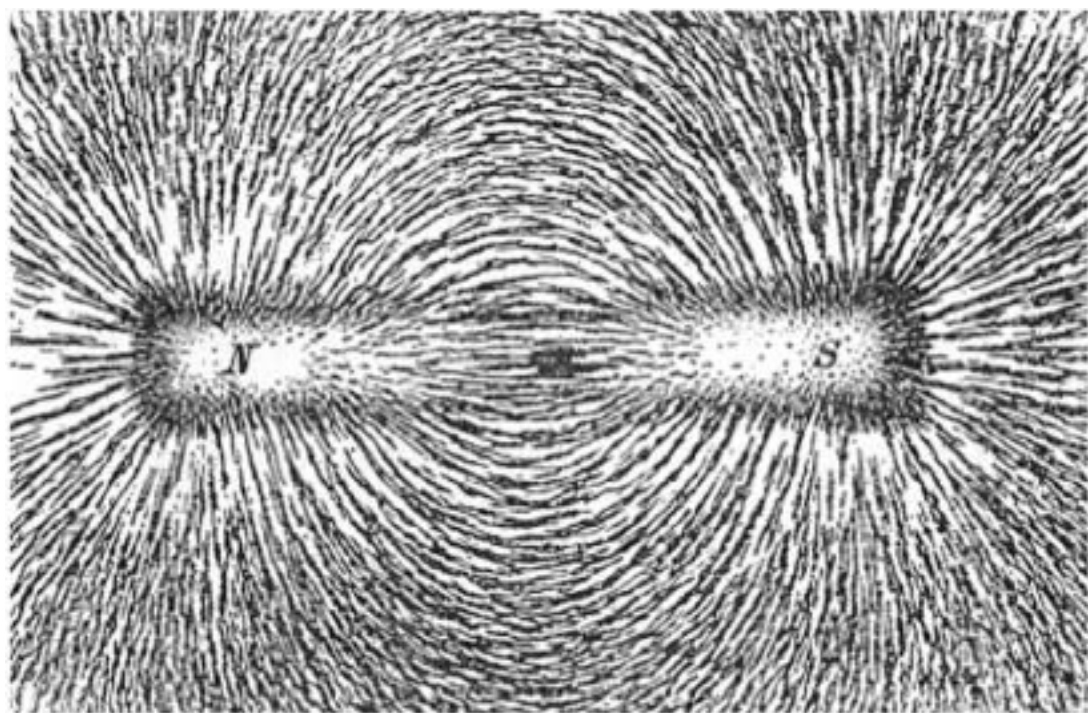


Fig.3. Limaduras de hierro dispuestas a lo largo de las líneas de fuerza del campo magnético generado por un imán. *Biangular Coordinates Redux: Discovering a New Kind of Geometry - Scientific Figure on ResearchGate*. Fuente:

https://www.researchgate.net/figure/Magnetic-lines-of-force-of-a-bar-magnet-shown-by-iron-filings-on-paper-11_fig4_233640347

Maxwell había leído los "*Experimental Researches on Electricity*" de Faraday, y había quedado convencido por las explicaciones e intuiciones de Faraday acerca de las líneas de fuerza. De modo que, a principios de la década de 1860, se propuso formular dichas ideas en términos matemáticos, una tarea que Faraday (cuyo conocimiento matemático, según se dice, era limitado) no había realizado. Es importante resaltar que, si bien Faraday no dotó de contenido matemático a sus descripciones de los campos eléctrico y magnético en términos de líneas de fuerza, demostró una gran intuición física al introducir el concepto de campo y aceptar su validez como realidad física subyacente a los fenómenos observados. Es por ello que también debe ser considerado un importante teórico, si bien esta faceta suya ha quedado a menudo oscurecida y poco reconocida, debido probablemente tanto a su innegable pericia en el plano experimental como a su falta de destreza matemática. "*Al principio me asusté cuando vi la cantidad de fuerza matemática relacionada con el tema*", le confesó Faraday a Maxwell [5] cuando este último introdujo por primera vez sus ideas en un artículo publicado en dos partes entre 1855 y 1856, titulado "*On Faraday's Lines of Force*".

Maxwell quería encontrar una teoría coherente que permitiera explicar de manera satisfactoria las diversas leyes formuladas de manera cuantitativa y precisa hasta la fecha por Coulomb, Ampère, Faraday..., que aparecían como piezas bien establecidas pero inconexas de un único puzle. Para ello, empleó el concepto de líneas de fuerza de Faraday -como ya se ha mencionado- y los conceptos introducidos por Poisson, Gauss y Green, pero además se sirvió de una analogía física: un modelo mecanicista del medio en el que se producían los fenómenos electromagnéticos (del éter en este caso, según el pensamiento de la época). Maxwell entendía este modelo como "*una ayuda al pensamiento*", en sus propias palabras [5].

Esta ayuda fue clave. El modelo de Maxwell fue evolucionando: empezó tomando elementos del flujo de calor y sus ecuaciones, y acabó representando los medios en los que se producen los fenómenos electromagnéticos como si estuvieran compuestos por esferas empaquetadas, entre las que se movían partículas pequeñas (comparadas con el tamaño de las esferas). El movimiento de estas partículas (corriente eléctrica) actuaba como cojinete y hacía girar por fricción las esferas, lo cual hacía que las esferas se achataran por los polos y se ensancharan por el ecuador, del mismo modo que le ocurre a la Tierra al girar. El ensanchamiento provocaba una fuerza efectiva radial en el plano del ecuador, y el achatamiento provocaba de igual modo una fuerza efectiva de 'succión' en la dirección perpendicular a este plano. La diferencia en la intensidad de los fenómenos en distintos medios (por ejemplo en el hierro comparado con otros materiales) se debía entonces a una distinta densidad de las esferas.

De esta forma, Maxwell logró explicar todos los fenómenos electromagnéticos conocidos hasta ese momento, y aunarlos todos bajo un conjunto unificado de ecuaciones. En otras palabras, puso las piezas conocidas del rompecabezas en su sitio. Pero no solo eso. Pudo explicar también lo que sucedía en la frontera entre dos materiales con propiedades eléctricas y magnéticas distintas, lo cual se desconocía. Más aún, Maxwell consideró que las esferas eran elásticas, y que por lo tanto se podían deformar, dando lugar a

pequeños desplazamientos de las partículas eléctricas y en consecuencia, a pequeñas corrientes (a este tipo de corriente, que aparece cuando el campo eléctrico varía en el tiempo, la denominó “*corriente de desplazamiento*”). Esto ocurriría en los materiales aislantes sometidos a campos eléctricos, mientras que en los conductores las partículas tendrían mayor libertad para poder moverse (esta característica de los materiales se explicaba por un distinto grado de elasticidad de las esferas). Éste era un fenómeno completamente nuevo, predicho acertadamente por Maxwell.

Maxwell se dio cuenta además de que las corrientes de desplazamiento generadas por la variación del campo eléctrico eran también fuente del campo magnético (es decir, producían un campo magnético). Al introducir esta novedosa contribución en las ecuaciones, pudo predecir que las perturbaciones (variaciones) en el campo eléctrico provocaban perturbaciones en la rotación de las celdas, y por lo tanto, en el campo magnético, y viceversa, y que estas perturbaciones se transmitían en forma de ondas transversales. Esto es justo lo que pasa en las ondas electromagnéticas. Maxwell calculó la velocidad de propagación de estas ondas, ¡y halló un resultado muy cercano a los medidos por el francés Armand Fizeau en 1849¹¹ y anteriormente por el gran astrónomo británico James Bradley en 1728! Acababa de descubrir y de predecir que la luz es una onda electromagnética¹². Sus resultados fueron publicados en un artículo en cuatro partes (1860-1862) titulado “*On Physical Lines of Force*”, y de esta forma completó la denominada síntesis electromagnética. El “modelo de vórtices moleculares”, como el propio Maxwell lo bautizó, resultó ser una de las ideas más provechosas de la Historia de la Ciencia.

La formulación matemática del Electromagnetismo como una teoría completa en términos de campos desarrollada por Maxwell supuso, y supone aún hoy en día, uno de los mayores logros intelectuales alcanzados por la humanidad. Asimismo, esta síntesis electromagnética supuso la unificación de los fenómenos ópticos y los fenómenos electromagnéticos dentro de una teoría coherente. Es importante destacar en este punto también a Olivier Heaviside, quien reformuló las ecuaciones de Maxwell y consiguió reducir su número de trece a cuatro, dotándolas así de su forma moderna (Fig.4).

¹¹El valor de esta medición fue mejorado entre los años 1850-1862 por el también francés Jean Léon Foucault. Estos experimentos para determinar la velocidad de la luz fueron sugeridos directamente por Arago a Fizeau y Foucault.

¹²La existencia de la relación entre la luz y el electromagnetismo había sido puesta de relieve por primera vez de manera convincente en 1845 por Faraday, tras numerosos esfuerzos infructuosos. De hecho, el efecto descubierto por Faraday también fue explicado por Maxwell en su artículo. Faraday descubrió que la dirección de vibración de un rayo de luz linealmente polarizado sufre una rotación cuando el rayo atraviesa un material colocado en el interior de un solenoide por el que circula corriente (un experimento muy similar había sido llevado a cabo en 1823 por William Herschel, quien sin embargo no logró observar el efecto buscado). Este descubrimiento sugería vívidamente que existía una relación estrecha entre ambos fenómenos. Además, el propio Faraday, antes de que Maxwell lo demostrara, también había pensado que los efectos eléctricos y magnéticos debían transmitirse a velocidades finitas, probablemente mediante vibraciones (ondas); así lo dejó escrito en una nota que hizo guardar en una caja fuerte en la Royal Institution en 1831.

Más de veinte años más tarde de la predicción hecha por Maxwell, entre 1887 y 1888, un físico alemán llamado Heinrich Hertz confirmaría experimentalmente la existencia de las ondas electromagnéticas. La idea de Hertz fue generar una onda que viajara hasta un receptor que consistía en un cable conductor en forma de

$$\nabla \cdot \vec{E} = \frac{\rho}{\epsilon_0}$$

$$\nabla \cdot \vec{B} = 0$$

$$\nabla \times \vec{E} = -\frac{\partial \vec{B}}{\partial t}$$

$$\nabla \times \vec{B} = \mu_0 \vec{J} + \mu_0 \epsilon_0 \frac{\partial \vec{E}}{\partial t}$$

Fig.4. Ecuaciones de Maxwell en el vacío, expresadas en su forma moderna. Fuente: elaboración propia del autor.

circunferencia con una pequeña abertura. Al atravesar la onda el circuito, el flujo magnético a través de éste variaría y se induciría en él una fuerza electromotriz, que se manifestaría por la producción de una chispa entre los extremos del cable. Hertz llevó a la práctica esta idea con éxito, completando de este modo la primera transmisión y recepción de ondas electromagnéticas, y confirmando también así las predicciones, y por lo tanto la teoría, de Maxwell. Hoy en día, la transmisión de ondas electromagnéticas es fundamental en el mundo moderno (piénsese en las telecomunicaciones, la radio, el wifi, etc.).

La Revolución Eléctrica.

En paralelo a los descubrimientos y los avances en la comprensión de la naturaleza del electromagnetismo, se produjo una auténtica revolución mediante la introducción de la bombilla, la corriente alterna, los generadores eléctricos y otros brillantes inventos que cambiaron el mundo para siempre. No es necesario enfatizar aquí en exceso la magnitud de la importancia de los inventos desarrollados a raíz del conocimiento generado por el descubrimiento de las leyes del electromagnetismo. Basta simplemente con imaginar por un momento cómo sería nuestra vida y cómo sería el mundo sin electricidad. La historia de todas estas invenciones es también fascinante, pero solo se abordará someramente debido a que no queremos desviarnos demasiado de la narración sobre el desentrañamiento de la naturaleza física profunda del electromagnetismo. Se mencionarán brevemente, eso sí, las importantes contribuciones hechas por Faraday, Swan, Edison y Tesla.

El primer descubrimiento de Faraday en el campo del electromagnetismo no fue la inducción electromagnética, sino la “*rotación electromagnética*” tal y como él mismo la bautizó. En 1821, Faraday colocó una barra imantada en el interior de un recipiente con mercurio y la lastró por uno de sus extremos, de manera que flotara verticalmente a modo de boya. Colocó también un cable conductor atravesando verticalmente el recipiente, e hizo pasar corriente por él. Como consecuencia, la barra comenzó a rotar alrededor del cable. Faraday había conseguido utilizar la corriente eléctrica para generar movimiento capaz

de efectuar trabajo: ¡había inventado el primer (y rudimentario) motor eléctrico! Las aplicaciones de los motores eléctricos hoy en día, a nivel industrial principalmente, son gigantescas.

Pero Faraday fue una persona prolija. En el mismo año 1831, diseñó el primer generador eléctrico, un disco de cobre cuyo borde se encontraba entre los polos de un imán en forma de herradura. Al girar el disco en el seno del campo magnético producido por el imán, se producía una fuerza sobre los electrones del mismo en la dirección radial, de manera que uniendo el centro y el borde del disco mediante un cable conductor y contactos deslizantes se generaba una corriente eléctrica en el circuito. De esta manera, se transformaba la energía cinética de rotación en corriente eléctrica; de una manera similar funcionan los generadores de las centrales que nos suministran energía, transformando la energía del movimiento de las turbinas en corriente, aunque en este caso la parte móvil suelen ser los imanes. Por supuesto, el funcionamiento del mundo moderno se asienta sobre la generación de energía eléctrica por estos medios.

También dependemos actualmente de la iluminación eléctrica. Como es bien sabido, Thomas Alva Edison patentó la bombilla en Estados Unidos en 1879. En realidad, la bombilla de incandescencia ya había sido inventada por Joseph Wilson Swan en torno a 1860, tras doce años de experimentos (hubo también otros inventores que desarrollaron diferentes modelos de bombilla). Sin embargo, estas bombillas no eran duraderas ni comercializables a gran escala. Edison probó más de 6000 materiales distintos hasta que encontró uno que le permitió obtener filamentos que duraran lo suficiente como para fabricar un artilugio que fuera comercialmente viable. El material escogido fue bambú carbonizado, que años más tarde sería reemplazado por el propio Edison por un filamento de tungsteno. Esta mejora convirtió a Edison en un hombre rico, y posibilitó la iluminación de hogares y ciudades, que comenzó a ser llevada a cabo al poco tiempo de producirse el descubrimiento. Por su parte, Swan se había vuelto a adelantar a Edison y había conseguido la patente británica en 1878, con un filamento de celulosa. En 1883, ambos fusionaron sus compañías en la “Edison and Swan Electric Light Company”, conocida comúnmente como “Ediswan”, zanjando de este modo la polémica sobre la prioridad en la invención de la bombilla ‘de uso práctico’.

Otra aportación clave al mundo fue la de Nikola Tesla, responsable de la introducción de la corriente alterna en 1888. La corriente alterna (que varía en el tiempo y revierte periódicamente su sentido) permite la transmisión y distribución eficaz de la energía eléctrica, evitando así las enormes pérdidas asociadas a la transmisión a grandes distancias de la corriente continua; hoy en día, las redes de alta tensión de distribución eléctrica funcionan con corriente alterna. Es famosa la disputa –denominada “la guerra de las corrientes”– que mantuvo al respecto Tesla con Edison, quien abogaba fuertemente (y de manera errónea) por el uso de corriente continua en la distribución de la electricidad. Tesla desarrolló además, en el mismo año 1888, un motor eléctrico capaz de funcionar con corriente alterna (en contraste con el primer motor diseñado por Faraday, que funcionaba con corriente continua), otra genial invención cuya enorme utilidad se extiende hasta nuestros días.

La Electrodinámica y la Teoría de la Relatividad Especial de Einstein.

Retomamos ahora de nuevo la narración sobre el estudio de los campos electromagnéticos: a principios del siglo XX, la teoría del Electromagnetismo clásico estaba firmemente establecida y había propiciado una revolución en la forma de vida de millones de personas mediante la introducción de sistemas de iluminación y la potenciación de la producción industrial. No estaba todo dicho respecto de la naturaleza última del electromagnetismo, sin embargo. El primer cuarto del siglo XX vio nacer dos teorías físicas, la Relatividad y la Mecánica Cuántica, que revolucionaron la Física, y eventualmente también, de nuevo, la forma de vida de millones de seres humanos mediante la invención del transistor, impulsor de la Revolución Electrónica. Ambas teorías aportaron nuevos conocimientos, más profundos, sobre el electromagnetismo.

El papel que jugó el Electromagnetismo, y en particular las ecuaciones de Maxwell, en el establecimiento de la Teoría Especial de la Relatividad fue absolutamente clave. El principio de relatividad de Einstein establece que las leyes de la Física son las mismas para todos los sistemas de referencia inerciales (sistemas en los que se cumple la 1ª ley de Newton, o ley de la inercia; estos sistemas se mueven con velocidad constante unos respecto de otros). Es decir, que las ecuaciones que determinan cómo se comportan los objetos cargados, por ejemplo, deben ser las mismas cuando se describe este comportamiento en un sistema inercial o en otro.

La pregunta entonces era cuáles eran las transformaciones (matemáticas) que permitían pasar de la descripción en un sistema inercial a la descripción en otro sistema inercial que se mueve con velocidad constante respecto al primero, manteniendo las leyes de la Física invariantes. Hendrik Antoon Lorentz había obtenido las transformaciones que llevan su nombre al imponer la invariancia de las ecuaciones de Maxwell frente a tales cambios entre sistemas inerciales. Estas transformaciones constituyen la base matemática de la Teoría de la Relatividad Especial, y eran conocidas por Albert Einstein cuando desarrolló esta teoría en 1905. Einstein demostró que las transformaciones de Lorentz describen de manera apropiada la Naturaleza, y que por lo tanto las ecuaciones de Maxwell son correctas; lo que había que modificar por tanto eran las leyes de la Mecánica de Newton (que no son invariantes), lo cual llevaba a predecir efectos verdaderamente sorprendentes (y anti-intuitivos) como la dilatación del tiempo o la contracción de las longitudes. Otra particularidad por cierto de las transformaciones de Lorentz deducidas a partir de las ecuaciones clásicas del Electromagnetismo de Maxwell, es que mantienen constante la velocidad de la luz.

Con la Teoría de la Relatividad Especial de Einstein, se puede resolver de manera clarificadora el siguiente problema: piense el lector en una carga que se mueve con respecto a un sistema de referencia (tres ejes de coordenadas "x", "y", "z"). En dicho sistema de referencia, el movimiento de la carga genera un campo magnético. Piense ahora en otra carga que se mueve en este campo magnético (en una dirección no paralela al mismo); la Física nos dice que dicha carga experimentará una fuerza. Ahora bien, piense ahora sobre

esta misma situación, pero vista desde el punto de vista de otro sistema de referencia en el cual una de las dos cargas está en reposo (velocidad cero). En este caso, un observador en dicho sistema debería concluir que no habrá fuerza (o bien porque la segunda carga no se mueve en el seno del campo magnético, o bien porque no hay tal campo magnético debido a que la primera carga no se mueve). ¿Entonces? ¿No decía el principio de relatividad que en todos los sistemas inerciales las leyes de la Física son las mismas? ¿Que si una carga experimenta una fuerza de Lorentz¹³ en un sistema inercial, también experimentará una fuerza en otro sistema inercial? La respuesta a esta aparente paradoja se encuentra en las transformaciones interdependientes de los campos eléctricos y magnéticos al pasar de un sistema a otro. En palabras del propio Einstein [6]:

“Lo que me condujo más o menos directamente a la Teoría Especial de la Relatividad fue la convicción de que la fuerza electromotriz que actúa sobre un cuerpo en movimiento en un campo magnético era nada menos que un campo eléctrico”.

De esta forma, queda puesta de relieve la más profunda e intrínseca interrelación entre los campos eléctricos y magnéticos, entre la electricidad y el magnetismo, cuya misteriosa conexión comenzó a destaparse con el experimento de Oersted a principios del siglo XIX. La otra gran teoría física moderna, la Mecánica Cuántica, también revela hechos profundos sobre la naturaleza de los campos electromagnéticos (la existencia del fotón y la naturaleza cuántica del magnetismo, principalmente)¹⁴. Y también surgió como solución a problemas directamente relacionados con el electromagnetismo (en particular, al problema de la llamada radiación del cuerpo negro”. Pero ésa es ya otra historia...

¹³La fuerza de Lorentz es la fuerza total (eléctrica y magnética) que experimenta una partícula cargada en el seno de un campo electromagnético. La expresión matemática para esta fuerza fue introducida por primera vez, no obstante, por Heaviside en 1889, y derivada posteriormente por Lorentz en 1895.

¹⁴Otro hecho sorprendente que explica la Mecánica Cuántica es la existencia a bajas temperaturas de materiales superconductores (como el aluminio o el niobio), en los que la corriente circula sin resistencia, y que hoy en día se utilizan por ejemplo para intentar fabricar ordenadores cuánticos. (Véase capítulo Computación Cuántica de Elías Fernández-Combarro).

Referencias:

- [1] Paul, C. y Valls, A. (2016). *El Magnetismo Rotatorio de Aragón*. 10.13140/RG.2.1.1694.6803.
- [2] Faraday, M. (1821-22). *Historical Sketch of Electro-magnetism*. *Annals of Philosophy* 1821: 195-200, 274-90; 1822, 19: 107-17.
- [3] García-Martín, D. (2018) *Michael Faraday, la inducción electromagnética y su enseñanza en 2º de Bachillerato*. *Ciencia y Educación* 6.
- [4] Parra Castillo, S. (2012). *Faraday. La inducción electromagnética*. RBA. España.
- [5] Sabadell, M.A. (2013). *Maxwell. La síntesis electromagnética*. RBA. España.
- [6] Resnick, R. (1977) *Introducción a la teoría especial de la relatividad*. Limusa. México.

Diego García-Martín

Junior research engineer en el Barcelona Supercomputing Center (BSC).

Realizando un doctorado en computación cuántica codirigido entre la Universidad Autónoma de Madrid (UAM) – Instituto de Física Teórica (IFT) y la Universidad de Barcelona (UB).

Capítulo 10

En la Prehistoria de la Astronáutica: Aerostación y Aviación.

Los primeros pasos de la Aeronáutica.

Prescindiendo de relatos mitológicos como el de Ícaro, se considera que el primer *vuelo* aceptablemente referenciado (que no documentado), punto de partida de la aeronáutica mundial, fue el protagonizado por el sabio hispanomusulmán Abdul-Kasim Abbas Ibn Firnas (Ronda, 810-Córdoba, 887), quien se supone que consiguió planear algunos metros desde La Arruzafa cordobesa sin matarse en el intento. También en nuestro país parece que tuvo lugar otra experiencia aeronáutica algo más prolongada (un *vuelo* de unos 300 metros planeando desde un altozano), protagonizada por el castellano Diego Marín Aguilera (Coruña del Conde, Burgos. 1757-1799) en el entorno de su villa natal, el 15 de mayo de 1793.

Pero en esas fechas de finales del siglo XVIII no serían dichas experiencias, intentando controlar la caída, las que fuesen a dar una solución científico-técnica al problema de la navegación aérea, sino las ascensiones en globos aerostáticos. De hecho, a los pocos días de los primeros vuelos realizados en Francia en 1783 (efectuados por François Pilâtre de Rozier y el marqués d'Arlandes con un globo de aire caliente, fabricado por los hermanos Montgolfier, el 21 de noviembre; y por Jacques Charles y Marie-Noël Robert, con un globo de hidrogeno, el 1 de diciembre), en la sesión de la *Academie des Sciences* de París, del 27 de diciembre



Fig.1. Ascensión de Pilâtre de Rozier y el marqués d'Arlandes el globo de los hermanos Montgolfier. (Archivo de ACC).

de 1783, Antoine de Lavoisier adelantaba el siguiente paso que debía dar la recién nacida aerostación: añadir a las envueltas llenas de hidrógeno elementos que permitieran dotarlas de dirección; es decir, pasar del *globo* al *dirigible*.

Mientras llegaba ese momento, los distintos países occidentales fueron incorporando a sus ejércitos esta nueva arma. Como es natural, los franceses

fueron los pioneros, y el 24 de noviembre de 1793, tras ordenar el *Comité de Salut Public* la construcción de globos para observación, a los físicos Jean-Marie-Joseph Coutelle y Nicolas-Jacques Conté, el 2 de abril de 1794 se decidía la creación de la *Compagnie des Aérostiers*, fijando su sede en Meudon (suroeste de París).

A lo largo del siglo XIX los globos evolucionaron hacia formas cilíndricas o ahusadas, en combinación con la aplicación de los primeros propulsores, destacando, entre las numerosas propuestas (las más de ellas disparatadas) de diferentes inventores, las pruebas de Henri Giffard, en 1852, con un primer *dirigible* propulsado por un motor de vapor; y la concepción, en 1872, por parte de Dupuy de Lôme, del *ballonet*, un pequeño globo de aire situado en el interior de su dirigible, que permitía mantener la forma de la envuelta (imprescindible para la estabilidad en ruta del aerostato) a pesar de las variaciones de presión de la atmósfera o de la temperatura del gas, solución que a partir de entonces se generalizaría en los futuros aeróstatos.

Pero la comunidad científica ha decidido que la Aerostación dirigida nació el 9 de agosto de 1884, cuando el dirigible "La France", de los ingenieros militares Charles Renard y Arthur Krebs, propulsado por un motor eléctrico a una velocidad de 6,5 m/s (casi 25 Km/h), describió una trayectoria cerrada desde Chalais Meudon retornando al punto de partida. Algunos años después, sería Alberto Santos Dumont el que continuase esta senda, en el ámbito civil, con sucesivos modelos *flexibles*, hasta lograr cierto reconocimiento público con su dirigible "Nº 6", con el que voló en torno a la Torre Eiffel el 6 de septiembre de 1901. Los globos empezaban a moverse por los cielos de forma dirigida, pero un problema seguía abierto: resolver su *estabilidad* de forma y en vuelo.

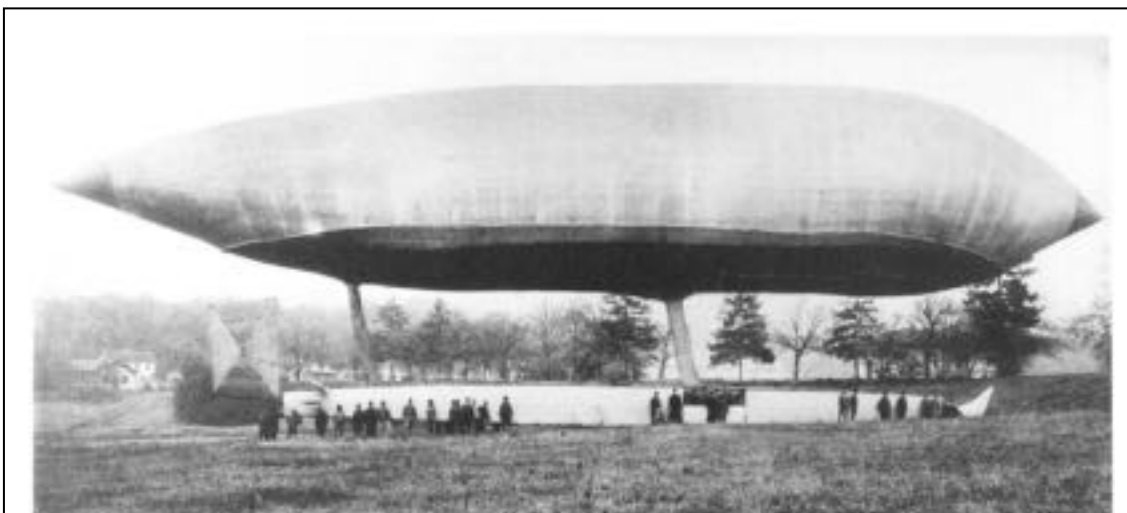


Fig.2. El dirigible "La France" de Charles Renard y Arthur Krebs. (Archivo de ACC).

En 1900 el conde Zeppelin siguió otro camino para resolver el problema de la estabilidad con su modelo de dirigible *rígido*. Este sistema optaba por usar una estructura de madera o metal, en cuyo interior varios globos impermeables contenían el hidrogeno conformando distintos compartimentos. La envuelta que recubría todo ello era de tela barnizada, sostenida

interiormente por la citada estructura que le daba la forma aerodinámica. Frente a los *flexibles*, los *rígidos*, junto al riesgo que suponía su fragilidad en caso de impacto, presentaban el inconveniente de la imposibilidad de plegar el globo una vez en tierra, aun vaciado de gas, lo que obligaba a edificar enormes hangares para cobijarlos.

En España suelen recordarse las experiencias aerostáticas (no tripuladas) de Agustín de Bethancourt y José de Viera y Clavijo con globos del tipo Montgolfier (inmortalizados en sendos óleos de Antonio Carnicero), a finales de 1783 en los jardines de Aranjuez y en los del palacio del marqués de Santa Cruz; así como los vuelos protagonizados por el italiano Vincenzo Lunardi, con globos de hidrógeno, desde El Retiro, el 12 de agosto de 1792, y desde el Palacio Real de Madrid, el 8 de enero de 1793. Sin embargo, en el ámbito de la investigación científica de los ilustrados españoles del entorno de Joseph L. Proust en Segovia deben destacarse los estudios que culminaron en los vuelos con un globo, también de hidrógeno, llevados a cabo por oficiales a las órdenes del conde de Aranda, en noviembre de 1792, desde el Real Colegio de Artillería y ante Carlos IV en El Escorial.

El Servicio de Aerostación Militar.

En cualquier caso, el primer paso hacia la institucionalización de la aerostación en nuestro país tuvo que esperar al 15 de diciembre de 1884, cuando se publicó, en *La Gaceta de Madrid*, el real decreto por el que se reorganizaba el Cuerpo de Ingenieros y se creaba la Dirección Técnica de Comunicaciones Militares, a cuyo cargo estarían el Servicio Telegráfico de Ingenieros y, en el seno de este, la 4.^a Compañía del Batallón de Telégrafos que “se ejercitará en la construcción e inflación de los globos aerostáticos y en su manejo, libre o cautivos, emprendiendo en la medida de los recursos de que puedan disponer, los ensayos y experiencias necesarios para las más útiles aplicaciones de estos nuevos instrumentos de guerra”.



Fig.3. Ascensión de globo-cometa Siegsfield-Parseval en Guadalajara. (IHCA).

Por real orden de 30 de septiembre de 1896, se creaba el *Servicio de Aerostación Militar*, independizado del Servicio de Telégrafos, con sede en Guadalajara y bajo el mando del comandante Pedro Vives Vich. El 12 de octubre de 1896 llegó el comandante Pedro Vives a Guadalajara, a tomar posesión de su nuevo puesto como jefe de la Compañía de Aerostación y establecer su sede en el cuartel de San Carlos. El primer globo esférico para

ascensiones libre comprado en Francia, "Marte", llegaría en septiembre de 1900; el segundo, "Venus", en enero de 1901. Entre ambos, en noviembre de 1900, se recepcionaría el globo cometa "Siegsfield-Parseval" para ascensiones cautivas.

Con el prólogo de la primera ascensión en vuelo libre, realizada por el propio Pedro Vives y el capitán Fernando Giménez Sáenz desde Guadalajara, el 11 de diciembre de 1900, las actividades del *Servicio* empezaría cuando, por real orden de 26 de junio de 1901, se suprimiese el Establecimiento Central de Ingenieros; el Servicio se convirtiera en Comandancia exenta de Ingenieros; la compañía pasase a denominarse "Compañía de Aerostación y Alumbrado de Campaña"; y, sobre todo, cuando comenzasen realmente las actividades del Servicio, la pionera Escuela Práctica, que proporcionaría las primeras promociones de aerosteros españoles: Alfredo Kindelán, Antonio Gordejuela, Emilio Herrera, etc.

La Escuela Práctica se convertiría en Escuela de Observadores en 1903, concebida esta como centro de formación del Ministerio de la Guerra para la ampliación de los conocimientos teóricos y prácticos de los oficiales de las diferentes armas del Ejército, no solo de los ingenieros. Por otro lado, tras la celebración en 1902 de la Conferencia de Aerostación Científica de Berlín, comenzaron en Guadalajara los trabajos de investigación

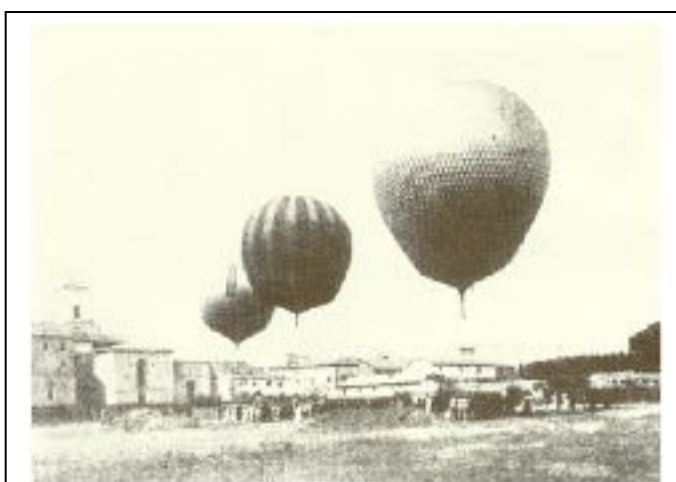


Fig.4. Los globos Júpiter, Marte y Urano en la observación del eclipse de sol en 1905. (Memorias de Emilio Herrera, 1988).

meteorológica, aprovechando las ascensiones tanto libres como cautivas. Esa iniciativa se consolidaría a partir del 3 de marzo de 1904, cuando se lanzase el primer globo sonda con instrumental para registrar datos de la atmósfera, y alcanzaría cierta notoriedad científica y social cuando, por designación de la *Comisión Internacional de Aerostación Científica*, el Servicio organizase en Burgos la observación en vuelo del eclipse de sol del 30 de agosto de 1905.

El Centro de Ensayos de Aeronáutica.

Sin embargo, las que pueden considerarse las páginas más significativas de la aerostación española, desde la perspectiva de la historia mundial de la aeronáutica, se empezaría a escribir con la aportación del ingeniero de caminos e inventor Leonardo Torres Quevedo.

En efecto, durante la primavera de 1901 Torres Quevedo comienza a concretar sus estudios sobre aerostación y entra en contacto con Edouard Surcouf, ingeniero director de los antiguos talleres aerostáticos de Gabriel Yon

en París. Un año más tarde, el 5 de mayo de 1902, D. Leonardo solicitaba la patente en Francia por “Perfectionnements aux aerostats dirigibles”, completada con una “Note sur le calcul d’un ballon dirigeable a quille et suspentes interieures”, presentadas ambas a las Academias de Ciencias de París y Madrid. Conjuntamente, la patente y la Nota proporcionaban un estudio en profundidad de todos los factores implicados en la solución del problema de la navegación aérea mediante dirigibles, especialmente de todas las fuerzas particulares que afectan la estabilidad de forma y en vuelo

Pero Torres Quevedo, en las investigaciones emprendidas al comenzar el siglo XX, había estudiado todos los posibles aspectos que podían llevarle a la solución del problema. En particular, había concebido un dispositivo para maniobrar a distancia los dirigibles sin poner en riesgo a los pilotos mediante el uso de las ondas hertzianas. Así, el 10 de diciembre de 1902 solicitó una nueva patente en Francia por un “Système dit Télékine pour commander à distance un mouvement mécanique”, en suma, su “telekino”, el primer mando a distancia de la historia y el precedente de los “drones” actuales.

Por real orden del Ministerio de Agricultura, Industria y Comercio, de 4 de enero de 1904, se creaban el *Centro de Ensayos de Aeronáutica* (dependiente de la Dirección General de Obras Públicas) y un laboratorio anejo, para la realización de los ensayos de los dirigibles patentados por Torres Quevedo y para dirigir la maniobra de motores a distancia con el “telekino”, nombrándose director a D. Leonardo.

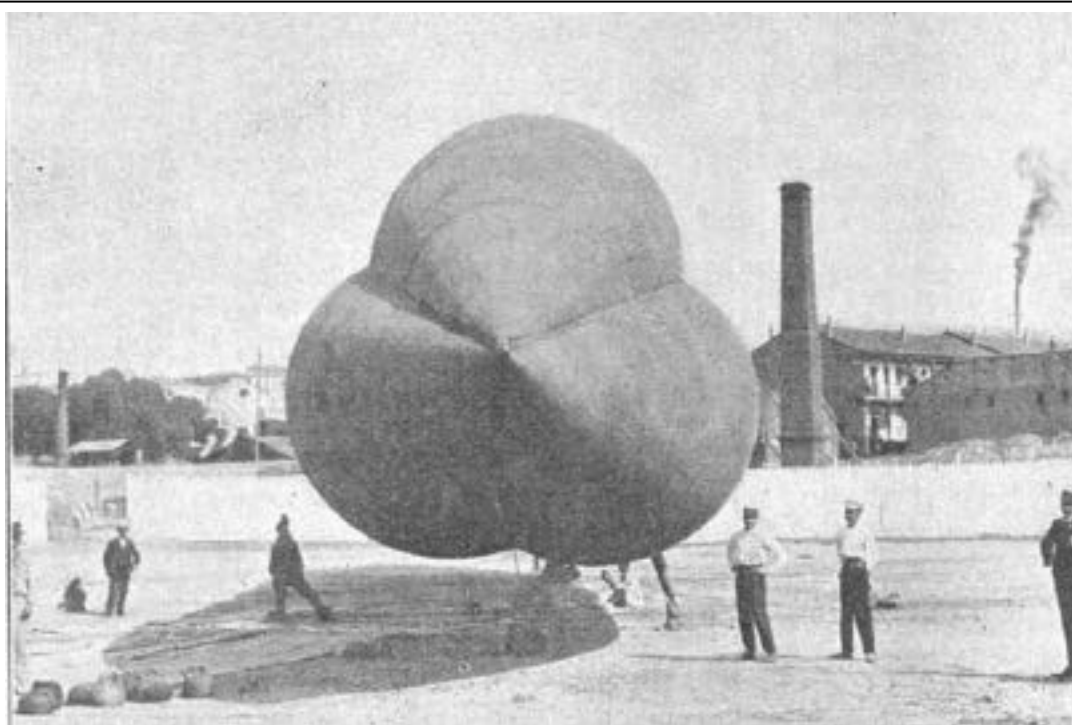


Fig.5. Dirigible ‘Torres Quevedo nº 1’ construido en el Frontón Beti-Jai de Madrid. (Archivo ACC).

En marzo de 1905 se incorporaba, como auxiliar técnico del Centro, Alfredo Kindelán Duany, cuyas tareas, durante la primavera de ese año, consistieron en la construcción de un pequeño generador de hidrógeno; de un aparato para comprobar la impermeabilidad de las telas; de dos modelos de dirigible a escala para realizar unos primeros ensayos de resistencia al viento y estabilidad de forma y en vuelo; la preparación de la viga funicular interior del dirigible y el despiece, preparación y cosido de las telas que debían conformar la envolvente. Finalizada la unión de la envolvente a la viga funicular, tras casi un año de trabajo, el 26 de junio de 1906, en las inmediaciones del *Gasómetro* de Madrid, se procedió al inflado del globo de 640 metros cúbicos con gas del alumbrado, operaciones que se realizaron con total normalidad y que culminaron con la solicitud de patente para “Un nuevo sistema de globos fusiformes deformables”. Éste reuniría las ventajas de los sistemas precedente (el *flexible*-francés- y el *rígido*-alemán-) y eliminaría sus desventajas: el dirigible sería flexible, desinflable, transportable y estable de forma por la presión interior. D. Leonardo había encontrado la solución real del “problema de la navegación aérea” mediante globos dirigibles *autorrígidos*.

Trasladado el globo a Guadalajara en julio, a lo largo de los meses siguientes Kindelán iría completando y acoplado al globo los elementos que permitirían hablar ya de un *dirigible*: barquilla, motores, cables de suspensión, timón de estabilización, etc., hasta llegar a la primavera de 1907, cuando puede decirse que ya existía realmente un aerostato completo.

Durante los días 12 y 13 de septiembre de ese año 1907 tuvieron lugar las pruebas del “Torres Quevedo nº 1”, pero no pudieron estudiarse ni la estabilidad de forma, ni su estabilidad y dirigibilidad en vuelo, por detectarse pérdidas de gas a través de la envolvente. Conservando la viga funicular



Fig.6. Dirigible ‘Torres Quevedo nº 2’ ensayado en Guadalajara. (Colección Francisco de la Torre).

triangular interior del dirigible primitivo, y aumentando el volumen del globo hasta los 960 m³, el 14 de junio de 1908 dieron comienzo las pruebas del “Torres Quevedo nº 2”. Por fin, España tenía su primer dirigible.

En Europa, los dirigibles de Zeppelin alternaban el éxito social inicial con la sucesión de accidentes; en Inglaterra fracasaban iniciativas como la del pretencioso “Nulli

Secundus”; en Alemania el prototipo de *Parseval* iniciaba sus ensayos; solamente Francia, con los semirrígidos *Lebaudy* y los flexibles *Astra* y *Zodiac* proponía unidades que parecían “aceptables” en 1908... pero que se verían superadas a partir de 1911 con el desarrollo de los “Astra-Torres”.

Aerostación y Aviación.

Realmente, durante los primeros años del siglo XX, los globos y los dirigibles siguieron constituyendo el núcleo de la Aeronáutica mundial. Y es que, aunque los hermanos Wright llevaban desde 1899 estudiando cómo volar con lo que acabaría siendo aeroplano (tarea en la que también estaban otros pioneros americanos como Samuel P. Langley), hasta el 17 de diciembre de 1903 no conseguirían planear unos 260 metros con el "Flyer", motorizado y catapultado en Kitty Hawk (Carolina del Norte, EE.UU.). En Europa, el primer vuelo lo realizaría Alberto Santos-Dumont el 12 de noviembre de 1906, aunque no sería hasta, precisamente, 1908, cuando pioneros como los hermanos Voisin, Henri Farman o los propios hermanos Wright trasladados a Francia, realizasen los primeros vuelos circulares controlados.



Fig.7. El Rey Alfonso XIII con Wilburn Wright en el aeródromo de Pau, Francia. (Revista Flight).

En España, a pesar del éxito de la colaboración entre la aerostación civil y la aerostación militar, al finalizar el verano de 1908 se produjo el desencuentro entre Torres Quevedo y los aerosteros militares y el *Centro de Ensayos de Aeronáutica* se veía obligado a dejar el *Parque Aerostático de Guadalajara*. En octubre de ese año Alfredo Kindelán y Emilio Herrera visitan el Polígono Militar de Aerostación de Tegel (Berlín), donde estudiaron el dirigible flexible tipo *Parseval*. A su vuelta a Madrid Kindelán declaraba que "el globo dirigible y el aeroplano no se excluyen: se complementan".

En España, a pesar del éxito de la colaboración entre la aerostación civil y la aerostación militar, al finalizar el verano de 1908 se produjo el desencuentro entre Torres Quevedo y los aerosteros militares y el *Centro de Ensayos de Aeronáutica* se veía obligado a dejar el *Parque Aerostático de Guadalajara*. En octubre de ese año Alfredo Kindelán y Emilio Herrera visitan el Polígono Militar de Aerostación de Tegel (Berlín), donde estudiaron el dirigible flexible tipo *Parseval*. A su vuelta a Madrid Kindelán declaraba que "el globo dirigible y el aeroplano no se excluyen: se complementan".

Aprobado el presupuesto del Ministerio de la Guerra, por real orden de 4 de enero de 1909, se concedió a Vives y a Kindelán una "comisión de dos meses a Inglaterra, Francia, Alemania e Italia para efectuar estudio de dirigibles y aeroplanos". A su vuelta emitían un informe aconsejando la compra de un dirigible (el "España") de la casa *Astra*, sociedad aeronáutica integrada en el conglomerado empresarial del magnate Henry Deutsch de la Meurthe, heredera de los *établissements* que dirigía Eduard Surcouf. Ambos oficiales consideraban que "sin ser un modelo perfecto, como no lo es ninguno de los actuales, es el que mejor conviene para nuestro país", por lo que el 21 de agosto de 1909 se firmaba la real orden autorizando su compra.

Por su parte, siendo imposible realizar nuevas pruebas en Madrid con el dirigible ensayado en Guadalajara, en abril comenzaba el traslado del *Centro de Ensayos de Aeronáutica* al hangar alquilado en Sartrouville a la casa *Astra*. Tras efectuar diversas modificaciones en el "Torres Quevedo nº 2", el 22 de octubre de 1909 el dirigible español efectuó un vuelo por los alrededores de

París. Aunque la imperfecta dirección del globo restó brillantez a la prueba, las ventajas del sistema torresquevediano quedaron manifiestas y el 18 de noviembre de ese año 1909 escribía Surcouf a Torres Quevedo para “confirmar que estaríamos dispuestos a construir, a nuestra costa, un modelo de dirigible basado en su principio”.

El 12 de febrero de 1910 se firmó el contrato de venta de las patentes francesa e inglesa del dirigible *autorrígido* a la sociedad *Astra*, que obtenía los derechos de explotación en todos los países del mundo menos España, donde quedaba fuera ante la posibilidad de que se construyeran dirigibles del sistema Torres Quevedo en el país que le vio nacer. Completando esta cláusula el inventor se comprometía con el Ministerio de Fomento español, el 27 de abril de 1910, a ceder todos sus derechos al Gobierno si este decidía construir algún dirigible de su sistema en España.

También de febrero de 1910 se realizaron pruebas en Pau (Francia) con el “España” fabricado por *Astra* para nuestro *Servicio de Aerostación Militar*; pruebas que terminarían accidentalmente el 23 de ese mes, desmontándose y trasladándose el dirigible a la capital alcarreña para completar los ensayos de recepción. Y, efectivamente, el 5 de mayo de 1910 el “España” efectuaba su primer vuelo entre Guadalajara y Madrid, aprovechando para pasar sobre el Palacio de Oriente con objeto de que, en palabras de Enrique de Aufrán en *España Automóvil*, fuera “presenciada su majestuosa marcha por toda la Familia Real”.

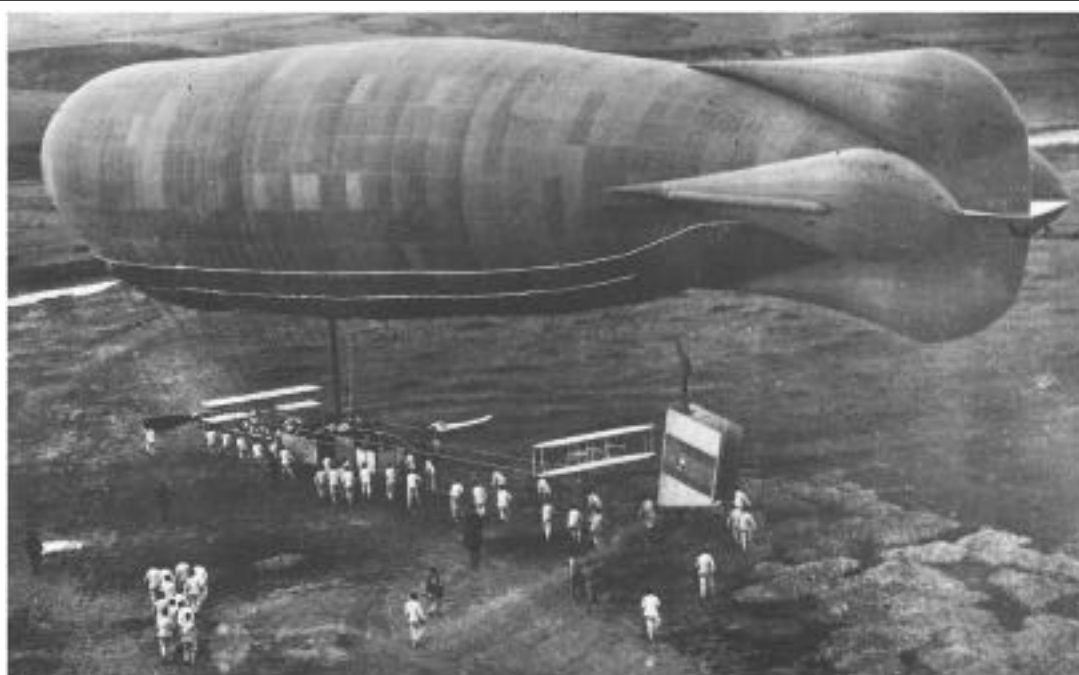


Fig.8. El dirigible “España” en Guadalajara. (IHCA).

Finalmente, el 8 de mayo se celebraba el acto de entrega solemne del dirigible al *Parque Aerostático*, por parte de la casa *Astra*, y se firmaba el acta de entrega. Nuestro ejército se conformaba con un deficiente dirigible francés,

que se caracterizaría por una lamentable hoja de servicios que solamente proporcionaría frustraciones y críticas, mientras había perdido la posibilidad de desarrollar en Guadalajara una industria aeronáutica en torno a un tipo de dirigible concebido por un ingeniero (civil) español, construido por un ingeniero (militar) español en instituciones españolas y ensayado con éxito en España.

Los comienzos de la aviación en España.

Incluso antes de su recepción, y constatado el error en la adquisición de un dirigible que inicialmente se pensaba que iba a poder intervenir en la Guerra de Marruecos, Pedro Vives dirigió las miradas hacia la naciente aviación. Por real orden de 2 de abril de 1910 se disponía: “Estando a punto de terminarse las pruebas de recepción del dirigible ‘España’ y siendo de la exclusiva competencia del Cuerpo de Ingenieros Militares cuanto se relaciona con los servicios de Aerostación, Aeronáutica y Aviación, el Rey (q.D.g.) ha tenido a bien disponer que por el *Parque Aerostático* se preceda al estudio del tipo de aeroplano que más convenga a nuestro Ejército”.

En realidad, existiendo el *Centro de Ensayos de Aeronáutica* (del Ministerio de Fomento) y una institución privada como el *Real Aeroclub de España*, carecía de sentido afirmar que era “de la exclusiva competencia del Cuerpo de Ingenieros Militares cuanto se relaciona con los servicios de Aerostación, Aeronáutica y Aviación”. Es más, antes de que los militares hubiesen mostrado interés por ello, ya habíamos tenido algunos pioneros españoles civiles participando en la naciente aviación. Así, el ingeniero de caminos Enrique Sanchís Tarazona había diseñado, a finales de 1908, un aeroplano siguiendo las ideas de los hermanos Wright, que terminó construyéndose durante los primeros meses de 1909 en las instalaciones de la casa *Astra* en Issy-les-Moulineaux. En esas mismas fechas el ingeniero industrial Gaspar Brunet Viadera emprendió la construcción de un biplano que, probado en Paterna (Valencia) en colaboración con el también ingeniero civil Juan Olivert, el 5 de septiembre de 1909, pasaría a la historia como el primer “vuelo” (un salto de unos pocos metros) de un aeroplano en España.

Pero el que puede considerarse el primer vuelo efectivo de un aviador español fue el de Antonio Fernández Santillana, constructor amateur afincado en Francia, quien se elevó por primera vez, el 8 de octubre de 1909, una distancia



Fig.9. Aeroplano, accidente y muerte de Antonio Fernández Santillana. (*l'Aerophile*, 1909).

de unos 200 metros en Issy-les-Moulineaux. Mayor éxito tendría Fernández el 5 de diciembre de ese año, cuando consiguió realizar un verdadero vuelo con retorno al punto de partida en el Certamen de Aviación de Niza. Lamentablemente, se estrellaría al día siguiente convirtiéndose en la primera víctima mortal de la Aviación española.

En cualquier caso, a lo largo de 1910 sí se fueron sucediendo los vuelos de aviadores franceses en España, comenzando con los de Julien Mamet con un "Blériot XI" en Barcelona, el 11 de febrero de 1910, y en Madrid el 23 de marzo; Léonce Garnier en Pamplona entre el 23 de marzo y el 15 de abril; o Hubert Le Blon, también con un "Blériot", en San Sebastián, donde se accidentaría y fallecería el 2 de abril de 1910; hasta llegar al 24 de noviembre, cuando Jean Mauvais aterrizó en Guadalajara con su aeroplano "Sommer", siendo recibido por Pedro Vives y los oficiales del *Servicio*, y agasajado por las autoridades alcarreñas.

Realmente, al crearse el 21 de septiembre de 1910 la *Comisión de Experiencias*, dependencia del Ministerio de la Guerra competente para regular la adquisición de los materiales del Cuerpo de Ingenieros, ya podía emprenderse la adquisición de aeroplanos para el *Servicio*, de modo que el 31 de octubre de 1910 salía nuevamente comisionado a Francia Alfredo Kindelán, con los fondos necesarios para comprar las primeras unidades.

Pero el *Polígono de Aerostación* de Guadalajara no podía acoger, simultáneamente, globos y aeroplanos, por lo que se buscaron diferentes ubicaciones donde instalar un nuevo aeródromo. Entre ellas destacaron dos: los terrenos de Los Cuatro Vientos en Carabanchel y la Dehesa de Santa Quiteria en Getafe, de entre los que se eligió el primero y a donde llegaron un sargento y seis soldados, el 16 de febrero de 1911, para comenzar la instalación de lo que sería el *Aeródromo Experimental de Cuatro Vientos*.

El 1 de marzo se recibieron los tres aeroplanos adquiridos por Kindelán, dos "Henry Farman" y un "Maurice Farman", y comenzó el montaje de los hangares tipo Bessonneau.



Fig.10. Pedro Vives con el instructor francés G. Osmont en Cuatro Vientos. (IHCA).

El 12 de marzo volaba desde Ciudad Lineal a Carabanchel el ingeniero industrial Benito Loygorri Pimentel, el primer español en obtener el título de piloto (en Francia), y el 15 de marzo empezaban las clases en la *Escuela de Aviación de Cuatro Vientos*, con el francés Geo Osmont como instructor y una primera promoción de oficiales del *Servicio* como alumnos: los capitanes Alfredo Kindelán, Emilio

Herrera y Enrique Arrillaga; y los tenientes Eduardo Barrón y José Ortiz Echagüe.

Construido un hangar tipo Vanniman en el aeródromo de Cuatro Vientos, después de tres vuelos desde Guadalajara el dirigible tuvo que ser desmontado en Carabanchel: la envuelta estaba inservible y había que cambiarla. En Francia, sin embargo, la casa *Astra* había comenzado los ensayos del “Astra-Torres nº 1” diseñado por Leonardo Torres Quevedo, encadenando éxitos: Premio Deperdussin en mayo; participación en el desfile militar de Longchamps, con motivo de la Fiesta de la República Francesa del 14 de julio; colaboración en las Maniobras del Este del Ejército Francés en noviembre, etc.

Además, en esos mismos años entraría en acción un nuevo ámbito aeronáutico hasta entonces ausente en España: el de los primeros constructores (artesanos) de aeroplanos. Entre ellos encontramos a Juan Vilanova y Luis de Acedo y su “Vilanova-Acedo” de 1911; José González Camó y, nuevamente, Luis de Acedo colaborando en la escuela privada de aviación del primero en Getafe, en 1912; José Barcala, Juan de la Cierva y Pablo Díaz construyendo su primer aeroplano “BCD-1” en 1912; Bernal y Mercadé en Barcelona; y Heraclio Alfaro en Vitoria; etc.

De la *Prehistoria* a la *Protohistoria* de la Astronáutica.

Al comenzar 1913 ya existían en España las tres ramas que puede considerarse que constituyen la *Prehistoria* de la Astronáutica: Aerostación (globos), Aeronáutica (dirigibles) y Aviación (aeroplanos). Al año siguientes se desencadenaría la Primera Guerra Mundial, en la que España no participó directamente, pero durante la cual una empresa española, la *Hispano-Suiza*, desde sus fábricas en Barcelona o mediante licencia en el extranjero, fabricaría miles de motores de aviación para Francia y el Reino Unido. Durante la Gran Guerra también operarían, al servicio de la *Royal Navy* y de la Aeronavale francesa, decenas de dirigibles *autorrigidos* trilobulados, construidos de acuerdo con las patentes de Leonardo Torres Quevedo.

Empezaba la transición de la *Prehistoria* a la *Protohistoria* de la Astronáutica, etapa que excedería ya los límites temporales de este capítulo.

Bibliografía:

- (1) André, P. 1994. *Le Temps des ballons. Art et Historie*. París, Éditions de La Martinière.
- (2) Atienza, E. 1994. *El general Herrera: aeronáutica, milicia y política en la España contemporánea*. Madrid, AENA.
- (3) Bengoechea, L. (dir.). 1988. *Historia de la Aviación española*. Madrid, IHCA.
- (4) Gomá Orduña, J. 1946. *Historia de la Aeronáutica Española*. Madrid, Ministerio del Aire.
- (5) González Cascón, A. 2001. Guadalajara 1896-1936, adelantada de la Aeronáutica española, en *Actas del VII Encuentro de Historiadores del Valle del Henares*, Guadalajara, pp. 463-488.
- (6) González-Granda Aguadé, R. 1994. *Crónicas Aeronáuticas*. 2 Vols. Madrid, IHCA.
- (7) González Redondo, F. A. y González de Posada, F. 2002. Ciencia aeronáutica y milicia. Leonardo Torres Quevedo y el Servicio de Aerostación Militar, 1902-1908. *Llull.*, vol. 25 (n.º 54), pp. 643-676.
- (8) González Redondo, F. A. 2003. El Centro de Ensayos de Aeronáutica: Ciencia, Técnica y Sociedad. En F. González de Posada (ed.), *Leonardo Torres Quevedo. Conmemoración del sesquicentenario de su nacimiento*. Madrid, Sociedad Estatal de Conmemoraciones Culturales, pp. 111-133.
- (9) González Redondo, F. A. 2009. *Protagonistas de la Aeronáutica. Leonardo Torres Quevedo*. Madrid, AENA.
- (10) González Redondo, F. A. 2010. Leonardo Torres Quevedo y el Servicio de Aerostación Militar, 1908-1914. Luces y sombras en la Aeronáutica española, *Llull.*, Vol. 33, pp. 57-88.
- (11) González Redondo, F. A. 2017. Los inicios de la aeronáutica española. En Asensio, S. (ed.) *La aportación de Manuel Rivera a la Navegación aérea*. Madrid, Ministerio de Economía-Ministerio de Defensa.
- (12) González Redondo, F. A. 2018. Los precursores de los servicios aéreos en España. En Utrilla, L. y Oller, J. *Historia del Transporte aéreo en España, 1919-1940*, pp. 1-26. Madrid, Fundación Enaire.
- (12) Lázaro Ávila, C. 1995. *La Aerostación Militar en España*. Madrid, Ministerio de Defensa.
- (13) Marimón Riera, L. 1979. *Historia de la Aeronáutica*. Madrid, Academia General del Aire.
- (14) Oller, J. 2011. *Protagonistas de la Aeronáutica. Pedro Vives Vich*. Madrid, AENA.
- (15) Salas Larrazábal, J. 1983. *De la tela al titanio*. Madrid, Espasa.
- (16) Utrilla, L. 2000. *Medio siglo de aeronáutica española*. Madrid, Fundación Aena.
- (17) Woolford, S. y Warner, C. 2015. *Flight: The Evolution of Aviation*. London, André Deutsch.

Francisco A. González Redondo

Licenciado en CC. Matemáticas (Universidad de Cantabria), Doctor en Filosofía (UCM) y Doctor en CC. Matemáticas (UPM).

Profesor Titular de Historia de la Ciencia en la UCM.

Capítulo 11

Tiempo y clima: desentrañando los secretos de la atmósfera terrestre.

Hablamos del “tiempo que hace” a diario con los familiares, con los amigos, con los vecinos. En ocasiones, equiparamos, erróneamente, condiciones “atmosféricas” puntuales con rasgos climáticos de un lugar y hablamos de “climatología adversa” que ha impedido el desarrollo de un evento deportivo, de una celebración, de un desfile. Viajamos a países con climas “exóticos” tan distintos a los que estamos acostumbrados en nuestra ciudad de residencia. Nos asombramos con el poder destructor de un huracán, de una borrasca “explosiva”, de una lluvia torrencial. Y valoramos positivamente el hecho de poder disfrutar de “buen tiempo”, con sol, calor y sin lluvias, en nuestro descanso estival.

Tiempo y clima se han convertido en elementos básicos de nuestro comportamiento diario, que condiciona nuestra actividad, nuestro ocio. Y llegamos a “exigir” una temperie a nuestra medida, a nuestra conveniencia...Y provocamos, con enorme irresponsabilidad, cambios en el clima que ha permitido la vida en la superficie terrestre. Y las disciplinas científicas que se encargan de su estudio –meteorología y climatología- tienen un pasado, un presente y un futuro, apasionantes; una historia llena de avatares que hoy nos permiten hacer pronósticos a unas horas o a una centuria.

¿Qué es el tiempo? ¿Qué es el clima?

De entrada, es importante que los conceptos “tiempo” y “clima” no son sinónimos, ni definen el mismo fenómeno. Tiempo es el estado de la atmósfera en un momento del día. Es cambiante, coyuntural. Y precisamente ahí, reside su “magia”, en su diversidad diaria. No hay un día con un tiempo igual a otro; puede ser semejante, pero nunca será igual. Siempre cambiará el valor de algún elemento atmosférico (medio grado de temperatura, un minuto menos de luz, un milímetro más de lluvia, etc.). El castellano tiene un vocablo precioso para referirse al tiempo atmosférico diario, la temperie. Por contra, el clima forma parte de la “estructura” de un territorio, es el ambiente permanente que se da en un espacio geográfico y que viene definido por los valores medios de los elementos atmosféricos básicos (temperatura, precipitación, viento, etc.), obtenidos durante un amplio período de tiempo; al menos, deben manejarse datos estadísticos de tres decenios para poder validar la pertenencia de un territorio a una variedad u otra de clima. El castellano clásico tiene un término precioso para referirse a las condiciones climáticas de un lugar, el temple. El climatólogo francés Pédelaborde, nos ofreció a mediados del pasado siglo, la

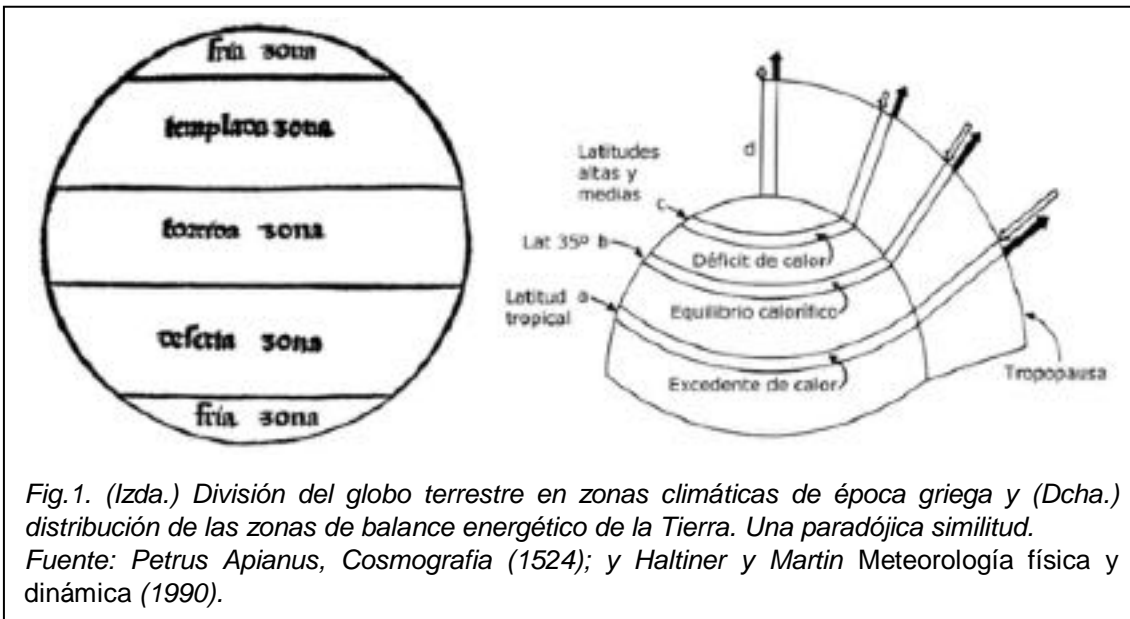
definición seguramente más completa y sencilla de concepto clima, al concebirlo con el conjunto de fenómenos atmosféricos “percibidos y vividos” por el ser humano a lo largo de su vida.

Una ciencia, con pasado milenario, y evolución acelerada en apenas dos siglos.

En dos siglos, los dos últimos, se han desentrañado los secretos de nuestra atmósfera. Bueno, decir esto puede ser algo pretencioso, porque todo avance científico tiene antecedentes donde asentarse y evolucionar, y en el caso de la meteorología esto arranca en época griega. Es entonces cuando se escribe el primer tratado de meteorología (s. V a.C.) en el que su autor, Aristóteles, intenta dar respuesta a los fenómenos atmosféricos conocidos entonces por el ser humano y también los no atmosféricos como el funcionamiento de los volcanes o de los terremotos, que eran considerados fenómenos originados por los “meteoros” (μετεωρος, lo que está en el aire). En el escrito de Aristóteles se contienen teorías y explicaciones que, en buena medida, han permanecido hasta la actualidad (papel protagonista del Sol) y otras que han sido rebatidas, afortunadamente, en aras al avance de la ciencia atmosférica. Pero tendrían que pasar algunos siglos para que comenzaran a desentrañarse los secretos del funcionamiento climático del planeta respecto a las ideas aristotélicas que defendieron eruditos romanos, como Plinio el viejo, y de época medieval (San Alberto Magno).

Y será precisamente la ampliación de los horizontes geográficos, el descubrimiento de nuevas tierras a partir del siglo XV, los que resultarán decisivos para el avance de la meteorología y climatología terrestre. Y en esto tienen un protagonismo destacado nuestros viajeros al Nuevo Mundo que luego dejaron escritos de sus experiencias en aquellas tierras. El jesuita José de Acosta, cronista de Indias, fue el primero que desterró científicamente la idea aristotélica de que en las tierras próximas al Ecuador no podía haber vida humana por el supuesto calor tórrido que allí se registraba. Nada mejor como la visita personal a las regiones geográficas en cuestión –ámbito intertropical- para comprobar que eso no es así, que en la línea ecuatorial y sus territorios próximos el clima tropical lluvioso y con temperaturas cálidas, pero no tórridas, permite la vida humana, como sabemos. En los escritos del padre Acosta¹⁵ descubrimos, también, que el aire se enrarece con la altitud ocasionando malestar y dolor de cabeza (soroche) cuando se asciende a una alta cumbre, y conocimos un fenómeno climático que ocupa portadas en los medios de comunicación en la actualidad, el famoso “Niño” que trastorna la circulación de la atmósfera algunos años dando lugar a lluvias donde no suele llover y a sequías donde lo hace habitualmente. Son avances fundamentales para las ciencias del tiempo atmosférico y del clima, cuyo contenido se perfeccionará en los siglos siguientes cuando se disponga de aparataje para realizar mediciones precisas de los elementos climáticos (temperatura, presión, humedad, lluvia, radiación solar, etc.).

¹⁵ Ver. ACOSTA, J. de, *Historia Natural y Moral de las Indias*, 1590. Ed. José Alcina Franch, Madrid, Ed. Historia 16. Crónicas de América nº 34, 1987, 515 p.



La importancia de disponer de aparatos de medida.

En efecto, toda ciencia que se precie requiere de un instrumental básico para obtener los datos que permiten proponer o comprobar hipótesis y establecer teorías, manejando, claro está, métodos de trabajo que permitan la universalidad de los resultados obtenidos. Así ocurrió también con la meteorología que fue completando la relación del aparataje básico para la observación y medición de las variables climáticas a lo largo de la Edad Moderna. Estamos en un contexto científico y filosófico que, además, hace una apuesta clara por la experimentación y la toma de datos como paso previo a la formulación de teorías (Galileo, Bacon, Hume, Descartes, Newton).

El siglo XVII es el siglo de la aparición de los aparatos de medida de los elementos atmosféricos, que se irán perfeccionando en los siglos posteriores. Galileo inventaría el primer instrumento capaz de medir los cambios de temperatura: el termoscopio neumático (1607). En 1639, Benedetto Castelli idea un primer pluviómetro que será perfeccionado en 1662 por Christopher Wren con un sistema de cubetas basculantes y mejorado en 1670 por Robert Hooke. En 1641, Fernando II, Gran Duque de Toscana, construye el termómetro de bulbo de alcohol con capilar sellado; el instrumento estaba provisto de un tubo de vidrio con alcohol marcado con 50 grados, pero no utilizó el cero como un punto fijo. Por su parte, en 1644, tras llevar a cabo su famoso experimento, Torricelli construye el primer barómetro de mercurio. El primer higrómetro nació de la inventiva del físico francés Guillaume Amontons que lo presentaría en 1687 en la Academia de Ciencias francesa. Por último, el primer anemómetro para la medida de la velocidad del viento lo construye en 1667 Robert Hooke. La puesta a punto de este instrumental meteorológico permitirá el desarrollo de los primeros embriones de redes meteorológicas. La primera de ellas, integrada por una decena de observatorios, fue instalada, a instancias del Gran Duque Fernando II de Toscana, por Luigi Antinori y funcionó entre 1654 y 1667. Será el germen de otras redes de observación que se irán ampliando y consolidando en los siguientes siglos y hoy son una

pieza fundamental para las predicciones meteorológicas y los estudios climáticos.

En el estudio de los elementos atmosféricos, la medida de la presión atmosférica fue, sin duda, uno de los problemas importantes que tuvo que abordar la ciencia en el siglo XVII. Desde la época de Aristóteles habían perdurado dos ideas erróneas y generalmente admitidas. La primera de ellas era que el aire no pesaba y la segunda que no existía el vacío. Galileo rechazará ambas cuestiones y demostrará, mediante una serie de experimentos, que el aire pesa. Años más tarde, en 1643, su discípulo Torricelli resolvería el enigma, llevando a cabo una serie de experimentos en su laboratorio, llenando de mercurio un tubo de 1 metro de largo, (cerrado por uno de los extremos). Invertió dicho tubo sobre una cubeta llena de mercurio, de inmediato la columna de mercurio descendía por término medio, hasta una altura de 76 centímetros (760 mm). Torricelli interpretó que a esa altura se producía la influencia de la presión atmosférica. La comprobación en campo de este hecho la realizó, unos años después, en 1648, Florin Périer, por encargo de su cuñado Blaise Pascal, al medir la altura de una columna de mercurio a tres altitudes diferentes, durante su ascenso al Puy de Dome. En este experimento se demostró que la presión atmosférica decrecía con el incremento de altitud, si bien ese descenso resultaba cada vez menos rápido. Por su parte, en 1656, el alemán Otto von Guericke demostraría la existencia del vacío con su célebre experimento de los hemisferios de Magdeburgo, señalando que la presión atmosférica equivalía a un peso muy considerable.



Fig.2. (Izda.) Experimentos de Blaise Pascal en el Puy de Dome (modificación de la presión atmosférica con la altitud) y (Dcha.) de Otto von Guericke (demostración de la existencia del vacío). Fuente: MeteoFrance y Experimenta Nova (1672).

Antes de que concluya el siglo XVIII se dispone de las cuatro escalas termométricas más difundidas, esto es, las de Fahrenheit (1724), con importante perduración aún en el ámbito anglosajón, Reamur (1730), Celsius y la escala centígrada, estas dos últimas tenidas erróneamente por equivalentes. Hay que recordar que Anders Celsius propuso en 1742 el uso de una escala donde se atribuía el valor 100° al punto de fusión del hielo y 0° al de ebullición del agua, es por tanto inversa a la denominada "centígrada". Esta última sería difundida en Suecia tres años después por Linneo, aunque ya había sido manejada en esta posición por Jean-Pierre Christin en Francia en 1743, en su

famoso “termómetro de Lyon” donde asignó el valor 0° al punto de fusión del hielo y 100° al de ebullición del agua.

El resultado de la puesta a punto del instrumental meteorológico favorecerá el desarrollo de observaciones atmosféricas sistemáticas durante la segunda mitad del siglo XVIII y, sobre todo, en el último cuarto de esa centuria, cuando se proponen los primeros embriones de redes o servicios meteorológicos. En la creación y consolidación de estas redes de observación meteorológica jugarán un papel destacado las Sociedades de Medicina, debido a la relación entre el desarrollo de enfermedades y las condiciones atmosféricas de un territorio. Esta relación encontrará argumento filosófico dentro del ambientalismo que impulsarán algunos ilustrados de la época.

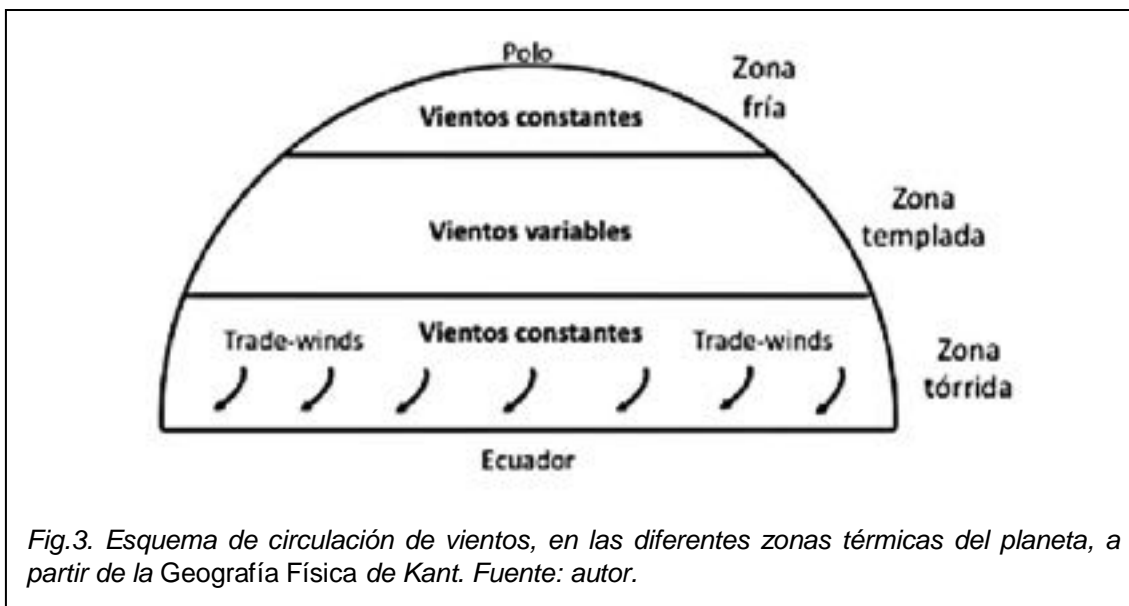
Una obsesión constante en la Edad Moderna: cómo soplan los vientos. La importantísima –y desconocida- aportación del filósofo-geógrafo Kant.

Junto a la aparición del instrumental meteorológico y de las primeras redes de recopilación sistemática de datos de los elementos atmosféricos principales (temperaturas, precipitación, presión, humedad, viento), los siglos XVII y XVIII supondrá un salto importante para el conocimiento de la circulación atmosférica y del reparto de los climas terrestre.

Un tema principal para la filosofía natural y la física durante este período fue desgranar el funcionamiento del viento. Al inicio de este período, las ideas de Aristóteles dominaban el pensamiento en esta cuestión. Para él el viento tiene por causa el Sol que mueve o transforma las “exhalaciones” seca y húmeda que existen en el aire. En dos siglos estas ideas experimentarán un vuelco radical y se irá conformando el corpus teórico de física del aire que resultará básico para la comprensión de la dinámica atmosférica. Diversos autores irán haciendo aportaciones para descifrar el movimiento de los vientos terrestres y, muy particularmente, de los vientos alisios, por su transcendencia para la navegación y el comercio entre Europa y los territorios conocidos y colonizados de África, América y Asia. Kepler, Bacon, Halley, Vareño, Lister y Hadley son los autores que realizan explicaciones para entender la circulación de los vientos en la superficie terrestre, sus trayectorias, su encorvamiento. La rotación terrestre comenzará a jugar un papel importante en esta explicación: los vientos “se retrasan” o “se adelantan” respecto a la velocidad lineal de rotación terrestre. En algún caso –Halley- se pensará que los vientos que se generan por el calor existente en el Ecuador soplarán, elevándose en la atmósfera terrestre, hasta el Polo y desde allí descenderán de nuevo al Ecuador generando un circuito único: la denominada “chimenea ecuatorial”. Esta teoría será desterrada años después por Hadley que precisará, con acierto, que ese circuito se interrumpe en latitudes subtropicales (hacia 30° de latitud), puesto que a partir de entonces y hasta el polo circulan vientos con otra componente, básicamente del oeste. Idea que se mantendrá hasta la actualidad, mejorando el conocimiento de aspectos regionales que inciden en la circulación de los vientos alisios.

En el tramo final de la edad Moderna, resultarán muy importantes y poco conocidas, las aportaciones a la ciencia climática de Immanuel Kant. Kant fue

un polímata, sin duda, uno de los mejores pensadores de la historia de la Humanidad; pero no solo fue filósofo. Fue también un geógrafo cuajado. O al menos eso traduce haber enseñado geografía durante cuarenta años en la universidad de Königsberg. Sus discípulos recopilaron sus apuntes en una Geografía Física, que reúne los conocimientos existentes en ese momento de la disciplina. Pero, además, en sus apuntes manuscritos, incluyó aportaciones novedosas para la explicación de hechos y fenómenos de la Naturaleza. Los capítulos que dedica a la explicación de fenómenos atmosféricos y hechos climáticos pueden considerarse como el primer manual editado de Climatología (1802). El viento y, en general, el movimiento del aire es el elemento climático más importante para Kant; describe las características de los vientos conocidos en su época y busca explicar sus causas. Y aporta un esquema novedoso de la circulación atmosférica en el globo terrestre. Según Kant “los vientos son más variables en la mitad entre un polo y el Ecuador. Tanto en la zona cálida y regiones adyacentes como en el cinturón frío y zonas vecinas son mucho más constantes”¹⁶. Señala que los cambios en el viento originados en una columna atmosférica, que pueden llegar a provocar “calmas y tormentas repentinamente o viento cambiante en los territorios bajos”. Este último aspecto tendrá gran interés porque inaugura la serie de estudios sobre el espesor de las capas atmosféricas, que constituirán referencia esencial para la navegación aerostática en el siglo XIX; aunque será ya en el siglo XX, con el desarrollo de la aviación y, posteriormente, de los satélites de finalidad meteorológica, cuando reciban completa y detallada explicación.



Dos siglos decisivos para el “descubrimiento” de la atmósfera terrestre.

El siglo XIX comienza con la propuesta de clasificación de las nubes que elaboró un farmacéutico inglés, Luke Howard. El fue el que dio nombre científico a las nubes en 1803, utilizando denominaciones en latín alusivas a su

¹⁶ Ver. Kant, I. *Physical Geography* [Rink's edition, 1802]. Vol. I, Introduction. §1, 2, 3, 4, y 5. (In. Watkins, E. (edit.) (2012) *Natural Science. The Cambridge edition of the Works of Immanuel Kant in translation*. San Diego, University of California, 818 p.

forma (stratus, empedrado; cirrus, mechón de pelo; cumulus, montón; nimbus, chaparrón). La clasificación de Howard tuvo eco muy favorable en los círculos científicos y culturales de la época. Entre 1821-22, el pintor inglés John Constable realizaría su “Estudio de nubes” basándose en la clasificación de Howard; por su parte, el genial polímata alemán J.W. Goethe manifestaría, en su “Ensayo sobre Meteorología”, las excelencias del sistema de Howard. Esta clasificación será la base del Atlas Internacional de Nubes de la Organización Meteorológica Mundial que se mantiene, con añadidos, hasta la actualidad. La última versión de este Atlas es del año 2017 y se puede consultar íntegramente en la red.

La descripción de los climas terrestres tuvo una aportación fundamental en los trabajos de A. de Humboldt, el noble prusiano de saber enciclopédico que viajó al Nuevo Mundo y nos dejó escritos extraordinario sobre su naturaleza y sus climas. En el inmenso haber científico de Humboldt destaca un hecho fundamental para la geografía ibérica: la verificación de la Meseta central española, descifrada a partir de las mediciones del barómetro en diferentes puntos de observación. Y también son destacables la invención y uso de unas isólineas planteadas por el naturalista alemán, necesarias para la representación cartográfica de los elementos del clima mundial como la isoterma y la isobara. El otro gran impulsor de la ciencia geográfica en la primera mitad del siglo XIX, el erudito alemán K. Ritter, destacará la importancia del clima en la distribución de las civilizaciones del mundo.



Fig.4. Primer mapa mundial de isotermas dibujado por A. von Humboldt en 1817.

Fuente: www.divulgameteo.es

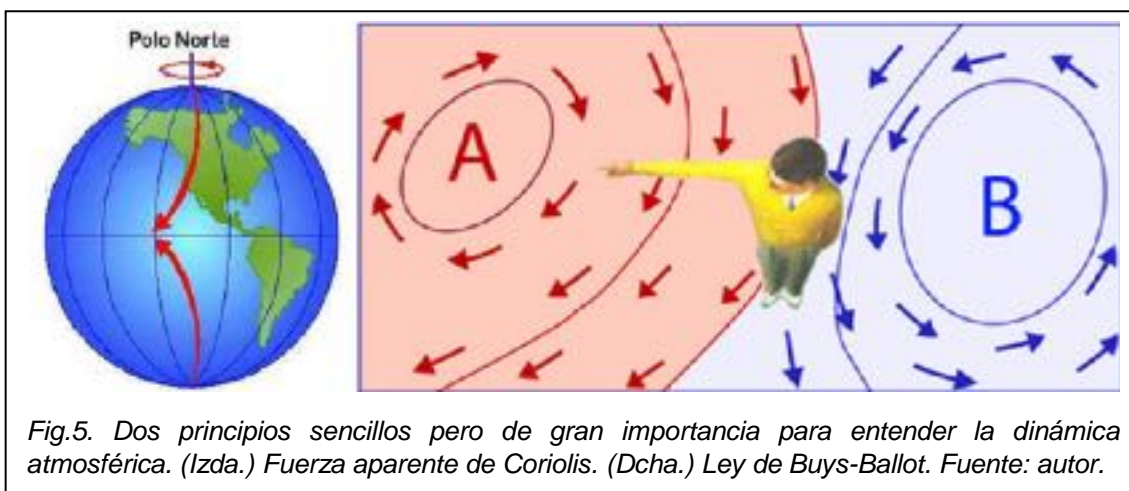
La dinámica de la atmósfera y la estructura interna de los centros de presión atmosférica (borrascas y anticiclones) son dos aspectos que avanzan de modo importante a lo largo del siglo XIX, en relación con el propio progreso de la física y las matemáticas. Y a ello se sumará el progresivo protagonismo que irá teniendo la circulación atmosférica en las capas altas que hoy se sabe esencial para los pronósticos meteorológicos. Se formularán, desde las primeras décadas de esta centuria, una serie de modelos explicativos del funcionamiento de las depresiones barométricas, que era lo que más interesaba en Europa central y septentrional de donde proceden los autores de las mismas.

El físico prusiano Dove enunciará su teoría cinética de las tempestades advirtiendo por vez primera que en el hemisferio norte el viento en las borrascas gira en sentido contrario al de las agujas del reloj y a la inversa en el hemisferio sur. Brandes, por su parte, dibujará en 1819 el primer mapa de distribución de presiones en superficie (mapa de tiempo) indicando las desviaciones respecto del valor medio, para lo cual utilizó las observaciones practicadas años atrás por la Sociedad Meteorológica Palatina. Este meteorólogo alemán enunciará su famosa teoría de circulación del viento en los sistemas de presión según la cual el viento sopla de los sectores de mayor presión a los de menos, siendo desviado hacia la derecha, en el Hemisferio norte, por la rotación terrestre. Otra contribución destacada de Brandes será su apreciación sobre el hecho de que las depresiones sobre el continente europeo se desplazan de oeste a este.

Dos leyes de la dinámica de fluidos, sencillas de formulación, permitirán comprender el movimiento del viento en relación con la rotación terrestre y con la presión atmosférica. La primera es la ley formulada por el científico francés Gaspard-Gustave Coriolis formulada en 1835, en realidad es la explicación de una fuerza dinámica afectada por la rotación terrestre –la denominada “fuerza de Coriolis”-. En realidad se trata de una fuerza ficticia que aparece cuando un cuerpo está en movimiento con respecto a un sistema en rotación. La fuerza de Coriolis siempre es perpendicular a la dirección del eje de rotación del sistema y a la dirección del movimiento del cuerpo vista desde el sistema en rotación. En síntesis, este principio se puede enunciar del siguiente modo: “Todo cuerpo que es impulsado o adquiere movimiento propio en un sistema rotatorio, como la superficie de la Tierra sufre una desviación lateral en sentido horario (hacia su derecha) en el hemisferio norte y en sentido antihorario (hacia su izquierda) en el hemisferio sur”. En síntesis, cualquier objeto en movimiento horizontal y que se desplaza sobre un sistema rotatorio, sufre una aceleración complementaria que lo desplaza lateralmente, en una dirección perpendicular a la dirección del movimiento del objeto. La comprobación de este hecho se ha hecho popular, e incluso reclamo turístico en localidades situadas en la línea ecuatorial, donde se muestra el giro que adquiere el agua que se escapa por la cañería de un lavabo, a uno u otro lado de la misma.

La segunda vendrá de la mano del meteorólogo holandés Buys-Ballot, quién en 1860 enunció su famosa ley de dinámica atmosférica: “Si una persona se coloca de espaldas al viento con los brazos en cruz, la presión atmosférica que soporta su mano izquierda es inferior a la que soporta su mano derecha”, por lo que las bajas presiones quedarán a su izquierda. En este contexto de

avances teórico-prácticos en el conocimiento de la dinámica atmosférica, unos años antes, el físico alemán Helmholtz había enunciado el principio de conservación de la energía y en 1864 el químico noruego Guldeberg formuló la ley de acción de masas de los equilibrios físico-químicos.



Desde Estados Unidos, Redfield estudiará los ciclones tropicales, de tanta trascendencia en este país; y Espy, dibujará mapas de tiempo indicando el rumbo de los vientos. Pero la figura más destacada será el marino Maury que en 1848 publicará una serie de cartas náuticas con indicación de los vientos dominantes en cada sector oceánico del mundo gracias a las cuales el viaje de ida y vuelta de Londres a Sidney se redujo a 160 días, cuando antes requería doscientos cincuenta.

La invención y difusión del telégrafo eléctrico en los países occidentales en las décadas centrales del siglo XIX permitió la transmisión rápida de datos meteorológicos, obtenidos en las redes de observación que van desarrollando diferentes países de Europa, América y Asia. En 1865 se estableció la Unión Telegráfica Internacional; el número de líneas y de mensajes transmitidos crecería exponencialmente a lo largo de la segunda mitad del siglo XIX y primeros años del XX. Ello supuso mejoras evidentes en la predicción meteorológica que podía disponer de datos de áreas geográficas diversas para la confección de mapas sinópticos y de boletines de pronóstico. Un acontecimiento bélico será determinante para el impulso de la creación de una red meteorológica entre países. El hundimiento a causa de una tempestad del navío de guerra Henri IV (13 de noviembre de 1854), que formaba parte de la flota anglo-franco-turca y piemontesa que operaba contra los rusos en la guerra de Crimea. La magnitud del desastre llevó a Napoleón III a preguntar al meteorólogo Le Verrier si se podría haber evitado de haberse conocido la llegada de dicha borrasca intensa, a lo que el meteorólogo francés respondió afirmativamente siempre y cuando se constituyera una red de vigilancia y aviso, mediante mensaje telegráfico, en los países europeos. Poco a poco, se consolidará la idea de crear un organismo internacional para fines meteorológicos, donde colaboraran todos los países ofreciendo sus datos y servicios de pronóstico. Esto ocurrirá en 1873 con la creación de la Organización Meteorológica Internacional, cuyo primer presidente será, justamente, el meteorólogo holandés Buys-Ballot. La OMI será el germen de la

actual Organización Meteorológica Mundial (OMM) organismo de las Naciones Unidas dedicado a las cuestiones del tiempo y clima mundiales.

Los últimos decenios del siglo XIX resultan muy activos en avances de las ciencias atmosféricas. Una serie de autores publicarán sus estudios sobre dinámica meteorológica, y sobre las diferentes capas que componen la atmósfera terrestre en relación con la mejora de los medios de navegación aérea (globos aerostáticos).

Destacarán los trabajos de Teiserenc de Bort que investigará con globos sonda sobre los niveles superiores de la atmósfera, y será descubridor de la “estratosfera”, además de reconocer la capacidad predictiva de la ciencia atmosférica en relación con los avances que se estaban produciendo a finales del siglo XIX. Hildebranson y Abercromby adaptarán la clasificación de nubes de Howard para convertirla en un catálogo que sería oficialmente adoptado por la recién creada Organización Meteorológica Internacional (1896). El alemán van Bebbber describirá en 1891 y sistematizará la trayectoria de las borrascas que desde el Atlántico penetran en el continente europeo en su desplazamiento hacia el este. Y el también alemán W. Koppen propondrá en 1903 una de las clasificaciones climáticas más utilizadas desde entonces, que manejaba valores medios de temperatura y precipitación para asignar con letras mayúsculas y minúsculas la adscripción de un observatorio a una variedad climática.

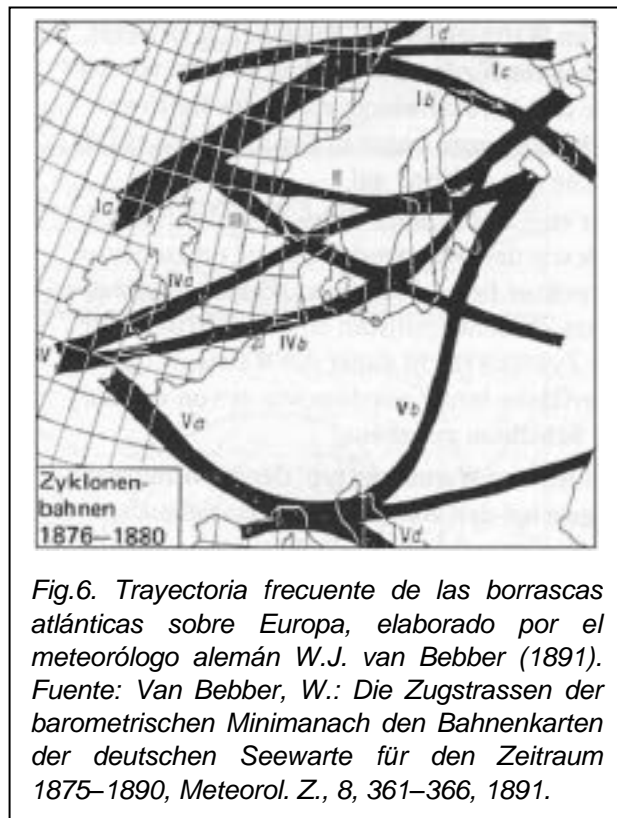


Fig.6. Trayectoria frecuente de las borrascas atlánticas sobre Europa, elaborado por el meteorólogo alemán W.J. van Bebbber (1891). Fuente: Van Bebbber, W.: Die Zugstrassen der barometrischen Minimanach den Bahnenkarten der deutschen Seewarte für den Zeitraum 1875-1890, Meteorol. Z., 8, 361-366, 1891.



Fig.7. Primer mapa de tiempo publicado en un periódico. The Times, 1 de abril de 1875. Fuente: <http://galton.org/meteorologist.html>

Y a ello se une el valor “social” de la meteorología, la popularización de la ciencia atmosférica, que vendrá de la mano de la publicación por el diario The Times del primer mapa de tiempo el 1 de abril de 1875. Dicha carta isobárica fue elaborada por Sir Francis Galton, psicólogo, antropólogo y geógrafo británico que había publicado, en 1863, su Meteorographica, primer ensayo sobre métodos de representación cartográfica del tiempo atmosférico, donde Galton propone una simbología propia para la elaboración de cartas sinópticas. A partir de ese momento,

diferentes periódicos europeos y estadounidenses incorporarían progresivamente una sección de información del tiempo atmosférico con reproducción de cartas isobáricas.

Los estudios de clima formarán una pieza importante de los ensayos que diversos autores, desde la geografía (Ratzel, Semple, Huntington), publicarán en las últimas décadas del siglo XIX y primeras del XX, en el contexto de la corriente filosófica del ambientalismo desarrollada a partir del evolucionismo de Darwin, destacando el papel del clima como elemento “condicionador” (determinismo climático) de la distribución de poblaciones y razas en el mundo. Estos escritos hay que interpretarlos en el contexto histórico en el que se publicaron; hoy rebasarían los límites de la racionalidad. No en vano, algunos de sus argumentos fueron utilizados por ideologías fascistas de la época para justificar la idea de la superioridad.

El siglo XX es un siglo de mejoras tecnológicas innegables que hoy son fundamentales para las ciencias atmosféricas (radiosondeos, aviación, satélites, computación). Pero es también una centuria de avances teórico-metodológicos necesarios para el desciframiento de la circulación atmosférica general y del funcionamiento de fenómenos meteorológicos (borrascas, ciclones, “corrientes en chorro”). En su primera mitad habrá una relación muy estrecha entre mejoras en el conocimiento de la atmósfera y su dinámica y la evolución del aparataje y la estrategia bélica (radar, “frente”). En su segunda parte vendrá condicionado por la formulación de la hipótesis de calentamiento térmico planetario y su progresiva comprobación que ocupa también estos primeros años del siglo XXI.

Comienza el siglo con un hito básico para la meteorología, la creación en 1917 de la Escuela Noruega de Bergen sin cuyo legado no se puede entender el avance contemporáneo de las ciencias atmosféricas. De este centro surgirán teorías y modelos de dinámica atmosférica esenciales: “frente polar” y proceso de ciclogénesis, estructura de las borrascas extratropicales, llamadas “noruegas” en su honor, distribución de las masas de aire en la superficie terrestre. Bjerkness, Bergeron, Solberg, Petterssen son algunos de sus miembros más destacados. Otros, como Rossby cobrará protagonismo en los años de la II Guerra Mundial y postguerra posterior, cuando descubra la clave de la circulación atmosférica en la alta atmósfera con su modelo ondulatorio (ondas de Rossby) que explica la dinámica en latitudes medias y altas como un continuo movimiento de masas de aire generado por corrientes de viento a alta velocidad (corriente en chorro o jet stream) y con efecto de la rotación terrestre. La denominación de “frente polar” propuesta por Bjerkness y Solberg, alusiva a la superficie de enfrentamiento entre una masa de aire polar y otra tropical sobre una superficie oceánica (océano Atlántico), alude a la fase bélica de la guerra de frentes que ocupó un largo período de la I Guerra Mundial.

Un nuevo y fundamental descubrimiento para las ciencias de la atmósfera tiene lugar en la Segunda Guerra Mundial en relación con la navegación aérea de guerra. Se descubre la existencia de corrientes de viento a gran velocidad en las capas altas de la atmósfera terrestre, a las que se bautizó con el nombre de “jet stream” (corriente en chorro), cuya circulación determina los movimientos de la atmósfera en las capas bajas, donde se

mueven las borrascas y los anticiclones, que vemos a diario en la información meteorológica de los medios de comunicación. En su descubrimiento, se disputan la autoría japoneses y norteamericanos que habrían encontrado estos vientos de alta atmósfera en los vuelos de los bombarderos de la guerra en el Pacífico. El descubrimiento de estas corrientes de viento a gran velocidad en la alta troposfera resultó fundamental no sólo para los estudios de meteorología y climatología, sino también para la navegación aérea transoceánica. El uso comercial de la corriente en chorro por la navegación aérea comenzó, el 18 de noviembre de 1952, cuando un vuelo de la compañía Pan Am entre Tokio y Honolulu se situó sobre la corriente en chorro polar a una altitud de 7.600 metros y pudo reducir el tiempo de viaje entre estas dos ciudades en más de un tercio, lo que significaba, asimismo, un importante ahorro de combustible.

Dos inventos de finalidad bélica también serán determinantes en el avance posterior de la meteorología. El radar (siglas de Radio Detection And Ranging) que debe mucho a las investigaciones de Nicola Tesla, fue creado en el Reino Unido en su modelo actual en 1935, sin olvidar algún antecedente alemán previo. Su empleo fue decisivo para el resultado del segundo conflicto bélico mundial, pero también para la ciencia meteorológica, puesto que ya en este período bélico, los operadores de radar notaban “ecos” de retorno de causados por tormentas. Será en la década de los años cincuenta del siglo XX, cuando la fuerza aérea de Estados Unidos comprobará su efectividad para la detección de fenómenos atmosféricos acuosos (lluvia, granizo, nubes de desarrollo...). Y a partir de 1964 se difundió su empleo para la predicción meteorológica en los países desarrollados, con modelos de respuesta doppler. Recordemos que el radar es un sensor “activo” que lanza un haz de ondas desde un punto fijo o móvil (avión) y espera la respuesta (eco) del objeto que interfiera en su trayectoria. En la actualidad los modelos mejorados de radar doppler son pieza básica para el “nowcasting” (predicción a corto plazo) meteorológico.

Algunos años después, en plena “guerra fría” y en el fragor de la carrera espacial, se desarrollará el segundo invento esencial para la meteorología y climatología modernas: el satélite de finalidad meteorológica. Recordemos que los soviéticos serían los primeros en lanzar un satélite al espacio (Sputnik-1) el 4 de octubre de 1957; pero serían los americanos, tres años después, el 1 de abril de 1960, los que pondrían en órbita durante poco más de dos meses, el primer satélite de finalidad exclusivamente meteorológica (Tiros-1). Desde entonces la tecnología satelital no ha dejado de mejorar y el programa de lanzamiento de satélites meteorológicos ha ido al alza en todo el mundo. Según su órbita se distinguen entre satélites de órbita polar, situados por término medio a 800 km. de la superficie terrestre, y satélites de órbita geoestacionaria, a 36.000 km. de la Tierra, situados en un punto fijo dentro del “anillo de Clark”, en honor al científico y divulgador Arthur C. Clark, autor del texto que dio origen a la película “2001, una odisea en el espacio”. Los polares tienen más resolución espacial, los segundos, más resolución temporal. Hoy en día no se entendería la predicción meteorológica y los estudios de clima regional sin la posibilidad de disponer de una imagen de satélite que, complementada con imágenes de radar, de sensores de actividad eléctrica y datos meteorológicos de estaciones de superficie, permiten ajustar los modelos

de pronóstico y mejorar el conocimiento de fenómenos meteorológicos singulares (tormentas convectivas, ciclones tropicales, tornados). En 1963 la Organización Meteorológica Mundial, heredera de la anterior Organización Meteorológica Internacional (vid. supra), puso en marcha su programa de Vigilancia Meteorológica Mundial, que es una pieza esencial de este organismo para el mejor conocimiento del tiempo y clima mundiales. En la actualidad está compuesto con varias unidades de satélite (polar y geoestacionario) y miles de puntos de observación en tierra y océanos.

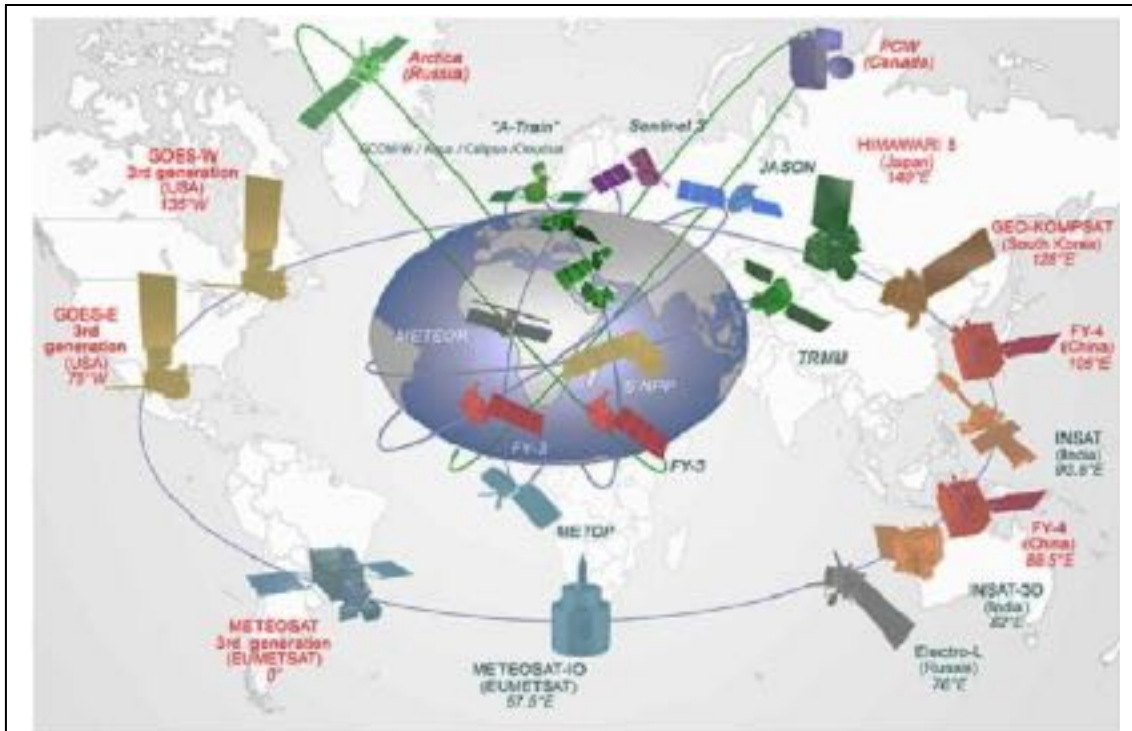


Fig.8. Programa mundial de Vigilancia Meteorológica. Satélites meteorológicos (2019). Fuente: OMM.

La aparición de los computadores, la mejora de los métodos matemáticos de tratamiento de datos y los avances en la dinámica de fluidos, en la química atmosférica (p.e. ozono estratosférico y troposférico) y en el mecanismo del balance energético del planeta han conducido a una mejora constante, “explosiva” por sus constantes progresos, de las ciencias atmosféricas en las últimas décadas del siglo XX. Se ha fijado el modelo de circulación atmosférica general que hoy se sabe tricelular (célula de Hadley, célula de Ferrel, célula polar) y condicionado por el movimiento de los vientos en la alta troposfera (corrientes en chorro ecuatorial, subtropical, polar y ártica). Se conoce casi con exactitud el balance energético del sistema Tierra-atmósfera, que resulta pieza clave del funcionamiento climático terrestre y de sus cambios. No en vano el actual proceso de “cambio climático” no es sino una alteración “humana” en dicho balance, puesto que la radiación de onda larga que emite la superficie terrestre y los océanos hacia el espacio exterior se queda confinada en los primeros kilómetros de la atmósfera terrestre al ser obligada a “rebotar” hacia el suelo por la acción de los “gases de efecto invernadero” de causa antrópica (CO₂, metano, etc.). Se ha avanzado, asimismo, en el conocimiento de un proceso atmosférico-oceánico de gran

impacto regional (el conocido como “Niño”, en realidad fenómeno ENSO) que aunque tiene su escenario de acción en la cuenca meridional del Pacífico, sus conexiones de circulación atmosférica, afectan a casi todo el mundo. Tras las primeras investigaciones sobre este fenómeno, relatado ya por el padre Acosta en el siglo XVI, como se ha señalado (vid. supra), a comienzos del siglo XX, fue en la década de los años sesenta del pasado siglo cuando se esbozaron las ideas principales de su funcionamiento y de la estrecha relación existente entre las dinámicas atmosféricas y oceánicas en dicha región del Pacífico. Este hecho ha sido clave para impulsar el conocimiento de otros mecanismos de oscilación atmosférico-oceánicas que existen en la superficie terrestre y oceánica de nuestro planeta, como la NAO, en el Atlántico Norte con importante efecto en el tiempo y clima de Europa.

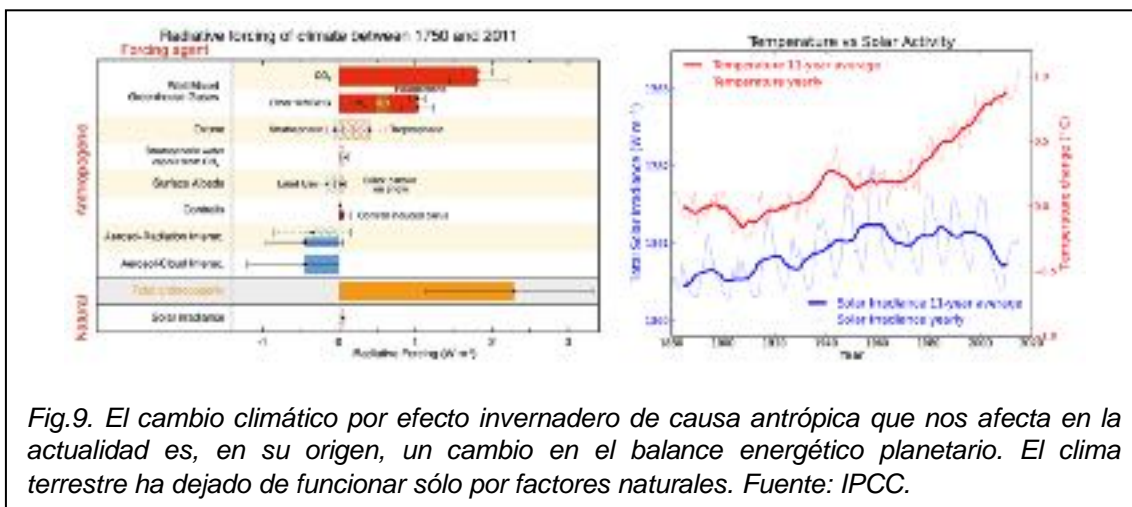


Fig.9. El cambio climático por efecto invernadero de causa antrópica que nos afecta en la actualidad es, en su origen, un cambio en el balance energético planetario. El clima terrestre ha dejado de funcionar sólo por factores naturales. Fuente: IPCC.

La precisión en los estudios de clima regional ha alcanzado niveles de excelencia, gracias al manejo de técnicas estadísticas y nuevos métodos de representación cartográfica (teledetección y Sistemas de Información Geográfica, SIGs). Se analizan series amplias de datos, mapas e imágenes para comprobar la evolución de las variables meteorológicas en un territorio (climatología analítica) o comportamientos de los sistemas de presión atmosférica (tipos de tiempo) y se avanzan comportamientos futuros (modelización).

Pero sin duda, el elemento que aglutina gran parte de los enfoques y de los avances en tiempo y clima es el referido “cambio climático” por calentamiento generado por causa antrópica. Esta hipótesis de trabajo, lanzada por varios investigadores en los años ochenta del pasado siglo, se ha convertido en un eje de investigación fundamental en todo el mundo. Y también en un argumento de políticas públicas, puesto que su mitigación depende, en buena medida, de la acción de los gobiernos en todo el mundo aprobando y cumpliendo acuerdos de reducción de emisiones de gases de efecto invernadero, causantes de la alteración del balance energético planetario y, en consecuencia del calentamiento en capas bajas de la atmósfera terrestre. Sin olvidar las acciones de adaptación a los efectos previstos de dicho calentamiento durante las próximas décadas que ocupan programas en las escalas estatales, regionales y locales. En 1988 se creó el Panel Intergubernamental para el Cambio Climático (IPCC), dependiente de las

Naciones Unidas, que es el organismo que tiene la responsabilidad de aprobar los informes de cambio climático en el mundo, que se elaboran regularmente. En la actualidad está en fase de elaboración el VI Informe del IPCC que tiene prevista su aprobación en 2021-22.

¿Y en el futuro?...reflexión final.

Como ha ocurrido en otras ciencias, el impulso contemporáneo en los conocimientos de la atmósfera terrestre ha sido intenso. Y lo sigue siendo. Pero este hecho no debe hacernos olvidar la historia de una ciencia milenaria de enorme influjo en el ser humano. Se ha pasado de responsabilizar a los dioses de los panteones mitológicos de las religiones antiguas (egipcia, sumeria, griega, romana) de los acontecimientos atmosféricos y climáticos, a disponer de una información detallada y puntual del estado del tiempo en nuestros teléfonos móviles. El futuro pasa por la inversión constante en investigación de la atmósfera que debe responder a una demanda social cada vez más exigente. Queremos saber el tiempo que va a hacer en mi ciudad, en mi barrio, con toda precisión. Y, en los próximos años, vamos a querer saber también con exactitud cuál va a ser el clima que tenga mi ciudad o mi barrio dentro de cuatro o cinco decenios. Seguramente ni una petición ni la otra podrá ser atendida al cien por cien, porque recordemos que la atmósfera terrestre es un sistema caótico cuyo conocimiento exhaustivo seguramente nunca alcanzaremos. Pero en doscientos años nos hemos ido aproximando a ese conocimiento de forma increíble.

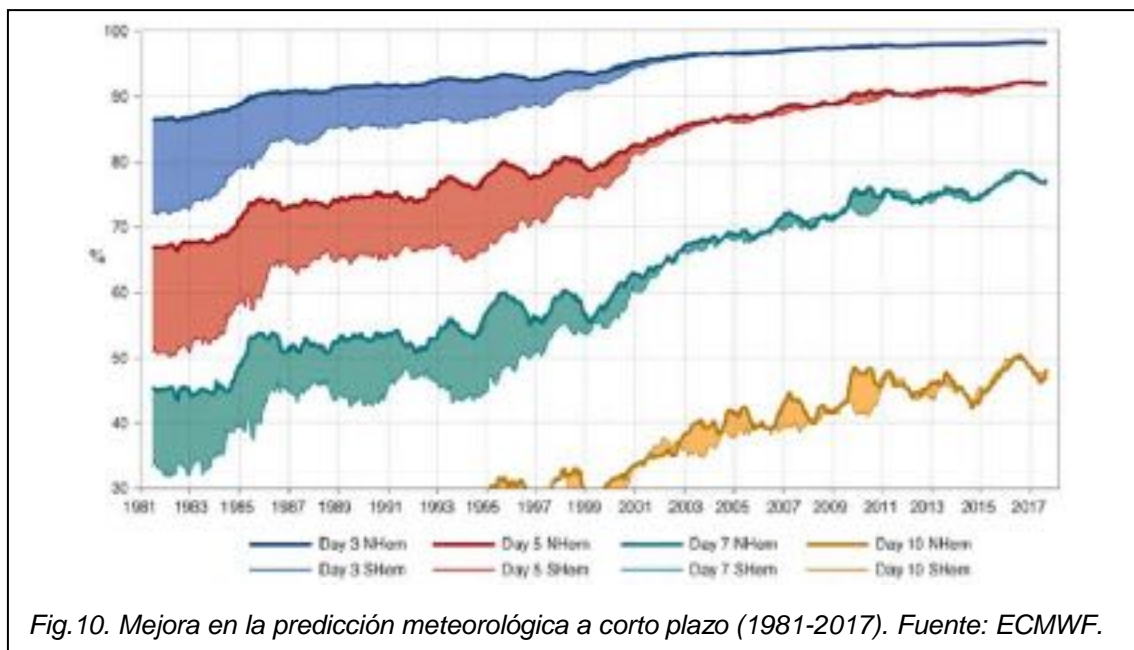


Fig. 10. Mejora en la predicción meteorológica a corto plazo (1981-2017). Fuente: ECMWF.

La mejora en el acierto de la predicción meteorológica es evidente. A pesar del habitual comentario coloquial sobre los fallos de pronóstico hoy en día con 24 horas de antelación el acierto en la predicción supera el 95%, sin ignorar que hay regiones del mundo de difícil pronóstico debido a su ubicación en relación con las circulación de vientos y sistemas de presión atmosférica y

sus singulares condiciones geográficas, como ocurre con el área del Mediterráneo. Lo mismo cabe decir de las mejoras en la modelización climática a medio y largo plazo. Desde los primeros informes del IPCC (1990) al último en vigor (2013-14) el afinamiento en las predicciones es evidente. Del 40% de fiabilidad estimada del primer informe del IPCC, al 85% del quinto Informe, vigente en la actualidad. Se estima que el próximo informe, el 6º informe del IPCC (2021-22), alcance una fiabilidad en su modelización del 90%.

En los próximos años, uno de los objetivos principales de las ciencias atmosféricas será dar respuesta a una aparente paradoja: la predicción meteorológica deberá mejorar en el largo plazo, es decir, los pronósticos a partir de 120 horas; y por su parte, los modelos climáticos tendrán que mejorar en el corto plazo, esto es, a una década vista para ofrecer proyecciones cada vez más realistas y espacialmente detalladas.

El conocimiento del tiempo y del clima terrestre va a seguir mejorando, sin duda. Pero siempre quedará ese punto de “magia”, incomprensible para el razonamiento humano que es esencia de la circulación del viento sobre la superficie terrestre y que sigue atrapando a aficionados e investigadores en la búsqueda de últimas respuestas sobre la dinámica de nuestra atmósfera.

Bibliografía:

- (1) CUADRAT, J. Mª y PITA LÓPEZ, Mª F. (1997) Climatología, Madrid, Ed. Cátedra, 496 pp.
- (2) BARRY, R.G. and CHORLEY, R. J. (2009): Atmosphere, Weather and Climate. (9ª ed.), London and New York, Ed. Routledge, 536 pp.
- (3) BURRIEL DE ORUETA, E. y OLCINA CANTOS, J. (2016) “Un período fundamental para la climatología española: el “descubrimiento” de la circulación atmosférica en altitud (1950 1980). Scripta Nova: revista electrónica de Geografía y Ciencias Sociales. Volumen XX. Número 525. 2016.
- (4) GIL OLCONA, A. y OLCINA CANTOS, J. (2017) Tratado de climatología. Alicante. Publicaciones de la Universidad de Alicante, 945 p.
- (5) Intergovernmental Panel on Climate Change, IPCC (2014): Climate Change 2013 and Climate Change 2014 (3 vols.) Disponible en: <http://www.ipcc.ch/> (Consultado: junio 2018.)
- (6) MARTÍN VIDE, J. y OLCINA CANTOS, J. (1996) Tiempos y climas mundiales. Climatología a través de mapas de tiempo e imágenes de satélite, Barcelona, Ed. Oikos-Tau, 308 pp.
- (7) MARTÍN VIDE, J. y OLCINA CANTOS, J. (2001) Climas y tiempos de España, Madrid, Alianza Editorial.
- (8) OLCINA CANTOS, J. (2014) “Referencias atmosféricas y avances para la ciencia climática en la obra de José de Acosta”. Scripta Nova. Revista Electrónica de Geografía y Ciencias Sociales [En línea]. Barcelona: Universidad de Barcelona, vol. XVIII, nº 478. <http://www.ub.es/geocrit/sn/sn-478.htm>.
- (9) PÉDELABORDE, P. (1958) Introduction a l'étude scientifique du climat, Paris, 150 pp.
- (10) SANTOS BURGUETE, C. (ed.) (2018) Física del caos en la predicción meteorológica. Ed. AEMET. Ministerio para la Transición Ecológica, Madrid, 1.116 p.
- (11) VIÑAS, J.M. (2016) El universo meteorológico. Un científico en las nubes. Ed. Materia- Descubrir la ciencia, Madrid, 145 p.

Jorge Olcina Cantos

Catedrático de Análisis Geográfico Regional.

Laboratorio de Climatología, Universidad de Alicante.

Capítulo 12

Santiago Ramón y Cajal: el nacimiento de la Neurociencia.

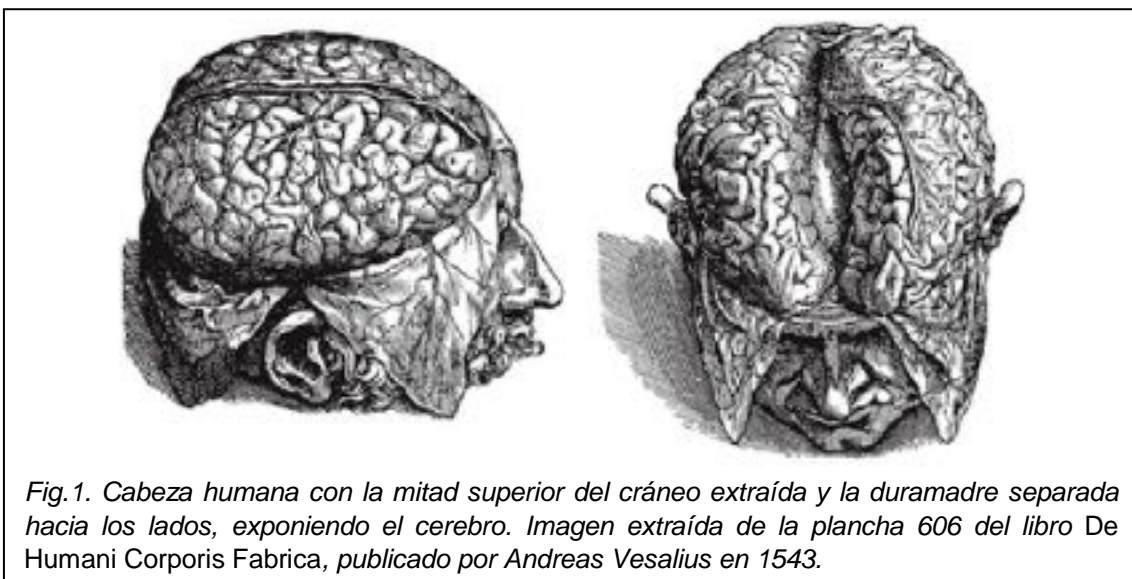
Siempre se ha dicho que podemos considerar a Cajal como el Padre de las Neurociencias modernas. ¿Quiere esto decir qué las Neurociencias comienzan con Cajal?

Veamos. Entendemos por Neurociencias al conjunto de todas aquellas disciplinas que abordan el estudio del Sistema Nervioso. Así, tenemos la primigenia Neurohistología o Neuroanatomía microscópica, y a cuyos cultivadores de la misma a menudo se les denomina morfólogos. Los primeros científicos o médicos o sabios, como se denominaban entonces a este tipo de estudiosos, que se aventuraban a escudriñar la anatomía microscópica de los seres vivos, lo hacían utilizando la Histología. Esta ciencia se desarrolla en el siglo XIX de manera paralela al perfeccionamiento del microscopio óptico, que empezaría a perfeccionar a final del siglo XVII, mediante el tallado de lentes, el comerciante de telas holandés Antoni van Leeuwenhoek (1632-1723). Con el paso del tiempo esta ciencia se vería completada por la Neuroelectrofisiología, que ya no basaba su estudio en la morfología de las células, sino en sus propiedades eléctricas. Podemos decir que el primer electrofisiólogo español fue Rafael Lorente de Nó (1902-1990), considerado el último discípulo directo de Cajal. Y tendrían que extinguirse Cajal y sus discípulos directos para que fueran apareciendo nuevas aproximaciones de estudio del Sistema Nervioso y se fuesen acuñando nuevos términos, como Neuroquímica, Neurofisiología, Neurofarmacología, Neurobiología molecular, del desarrollo, de sistemas, etc. Todas estas disciplinas que, como hemos dicho, estudian el Sistema Nervioso, se engloban dentro del término Neurociencias. Y vamos a poner un punto aquí, porque en la actualidad todo es neuro-algo. Parece ser que el cerebro o más propiamente, el término “neuro”, se ha puesto de moda y se usa, no para denominar una aproximación específica al estudio del cerebro, sino para fines meramente comerciales. Así, tenemos la Neuroeducación, la Neuroeconomía, la Neurogastronomía, etc. Términos a mi entender espurios y que no se deberían usar, pues utilizan el prefijo neuro- para potenciar otros términos bien establecidos, que no necesitan de ayudas extras.

Dicho esto, podemos decir que, Cajal efectivamente puede considerarse como el padre de las Neurociencias modernas, ya que posibilitó, con sus y acertadas interpretaciones, el estudio en profundidad del Sistema Nervioso que fue demandando poco a poco nuevas disciplinas de estudio. Pero veamos cómo se va desarrollando el conocimiento de este sistema. Miremos retrospectivamente y adentrémonos en la historia, en este caso, en la historia de la ciencia y en particular, en el descubrimiento de la célula nerviosa o neurona.

Desde luego, a principios del siglo XIX no se sabía prácticamente nada de la morfología de la célula nerviosa y, sin embargo, a finales de este mismo siglo ya se tenía la certeza de la existencia de la neurona y el mundo científico estaba familiarizado con la *Teoría Neuronal* que, enunciada por Cajal, ponía de manifiesto la existencia e individualidad de la célula nerviosa. Este dato fue importante pues cerraba la polémica abierta en torno a la *Teoría Celular*. Recordemos que con el descubrimiento de la célula en 1665 por Robert Hooke (1635-1703), las investigaciones del botánico Matthias Schleiden (1804-1881) y del zoólogo Theodor Schwann (1810-1882) concluyen, 200 años después, que la célula es la unidad estructural y funcional de todos los seres vivos (vegetales o animales). Y esta conclusión es apoyada por numerosos investigadores, entre los que cabe destacar a Robert Remak (1815-1865), Rudolf Virchow (1821-1902) y Albert Kölliker (1817-1905), dado que demostraron que las células se originan siempre a partir de otras preexistentes. De esta manera, es fácil imaginarse cómo los seres pluricelulares estamos formados por un montón de células de distintas clases que se agrupan selectivamente para formar los distintos tejidos que conforman los órganos, y éstos los seres vivos. Y cada uno de los cuatro tejidos básicos existentes, está formado por un tipo de célula. Así, el tejido epitelial está formado por la unión de las células epiteliales; el tejido conjuntivo por una buena variedad de células, dependiendo de su especialización, pero siendo la más representativa el fibroblasto; el tejido muscular por la célula muscular (miocito) y el tejido nervioso... por la célula nerviosa, claro, pero fue la última en descubrirse, por lo que la teoría celular anduvo mucho tiempo coja.

Si comenzamos esta historia en el siglo XIX, podemos decir que a principios del mismo se conocía bastante bien la estructura macroscópica de cerebro, cerebelo y médula espinal. Esto fue gracias a que, desde muchos años antes, existían buenos disectores de cadáveres que publicaban sus trabajos dando a conocer como estaba formado el cuerpo humano. Uno de los más famosos fue el belga Andreas Vesalius (1514-1564) que nos ha dejado su obra magna, "De Humani Corporis Fabrica" (Sobre la estructura del cuerpo humano), publicada en 1543 y de la que hemos entresacado la imagen mostrada en la Fig. 1.



Lo que no se conocía era la composición microscópica del tejido nervioso. Y esto era debido a varias razones. Una de ellas es que este material, tan excesivamente graso (hasta el 60% de su composición), era muy refractario a la impregnación. Es decir, no se conocían compuestos que fuesen capaces de impregnar las células nerviosas en su totalidad. Las impregnaciones parciales que se conseguían, gracias a tintes con alta afinidad por la cromatina, dejaban entrever numerosos cuerpos celulares de tamaños y formas diversas y que, además, marcaban el inicio de ramificaciones que, en número variable, emergían de los distintos cuerpos celulares. La complejidad de este sistema estaba servida. ¿Cómo estas células eran capaces de generar impulsos eléctricos y estos ser propagados para que saliesen del cerebro y conducidos a través de los haces de nervios? En aquel tiempo no se pensaban cosas raras; esto estaba asumido, porque unos años antes, en 1791, Luigi Galvani (1737-1798) publicó su famoso trabajo sobre la estimulación eléctrica en los nervios de las ranas, donde proponía que la contracción muscular está generada por corrientes eléctricas. Y como no se podía ver como estaba organizado el sistema nervioso, el tiempo fue pasando sin avanzar el conocimiento en esta materia. Sin embargo, los anatomistas de la época lo intentaban, aunque sin mucho éxito. Fue en 1863, cuando el alemán Otto Friedrich Karl Deiters (1834-1863), estudiando el núcleo vestibular lateral del cerebro, describe la existencia de un proceso tubular sin ramificaciones que se extiende desde el cuerpo de las células. Este proceso, que denomina cilindro-eje, hoy día es conocido como axón (término acuñado por Kölliker en 1896), y lo tienen casi todas las neuronas. A las demás prolongaciones de los cuerpos celulares los denomina procesos protoplásmicos, hoy conocidos como dendritas (término acuñado por His en 1889). Las descripciones de estos procesos celulares fueron pintadas y reproducidas en sus publicaciones, constituyendo una de las primeras imágenes de células nerviosas aparentemente teñidas en su totalidad. Por cierto, la técnica que desarrollo para sus estudios era harto complicada. Denominada como “disociación mecánica” consistía en aislar las células, utilizando unas agujas muy finas. En la Fig. 2 vemos uno de sus espectaculares dibujos.

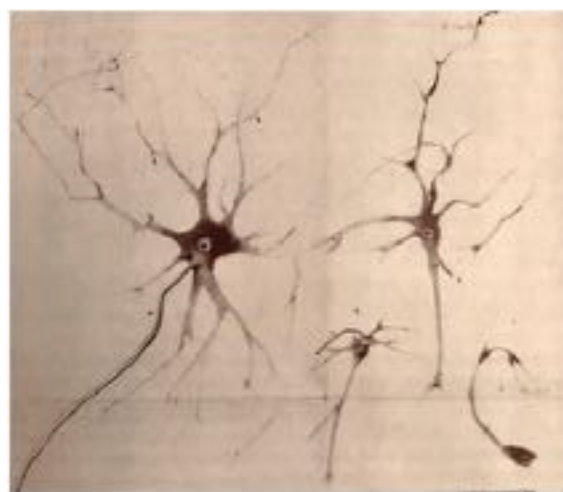


Fig.2. Diversas células ganglionares aisladas por Deiters de la sustancia gris de la médula espinal de buey, mostrando los procesos que salen del cuerpo celular. Técnica de la “disociación mecánica”.

Parece ser que Deiters, para sus estudios, fijaba las piezas en soluciones diluidas de dicromato potásico y las teñía con carmín. Posteriormente, como hemos mencionado, hacía disociaciones mecánicas, lo que le permitió describir los dos tipos de procesos de las células nerviosas. Hombre extremadamente hábil que, desgraciadamente, falleció de tifus el mismo año de estos descubrimientos, en 1863, contando con tan solo 29 años.

Y como acabamos de ver, ya hemos revelado una de las piezas clave en el conocimiento de la célula nerviosa. La tinción de las piezas con carmín, lo que le valió a Deiters a descubrir los procesos celulares. Sin embargo, este colorante no fue introducido por él, sino por el alemán Joshep von Gerlach (1820-1896), que en 1854 empezó a utilizarlo para estudiar el sistema nervioso. Tenía una serie de frascos sucios, a los que no había lavado después de contener una solución concentrada de carmín. Los enjuagó con agua y esta tomó una coloración rosa que aparentemente le llamó la atención. Como si fuese un juego, introdujo en uno de estos frascos una pieza fijada de cerebelo y se fue a casa. A la mañana siguiente observó la pieza y notó que la sustancia blanca de las circunvoluciones cerebelosas estaba inalterada, pero las capas de la sustancia gris de la pieza habían tomado de una coloración roja; mucho más oscura que la solución donde habían pasado la noche. Intrigado lo estudió al microscopio y vio que algunos gránulos de los cuerpos celulares se habían teñido de un color rojo oscuro. Como los procesos celulares y el fondo del tejido no se habían coloreado, los cuerpos celulares resaltaban en la preparación. Y esto era una peculiaridad de la mayoría de los colorantes, que se unían a estructuras celulares específicas. Por eso, se empezaron a combinar distintos colorantes para conseguir mejores impregnaciones. Uno de los más utilizados fue la hematoxilina.

En cualquier caso, así estaban las cosas; no se podía extraer demasiada información como para aventurarse a lanzar teorías arriesgadas. Sin embargo, en 1872, Gerlach aprovecha los datos anatómicos publicados por Deiters para postular que el cerebro estaba formado por una red protoplásmica, donde todas las células se anastomosaban (se unían) entre sí por los procesos protoplásmicos para formar una red difusa por donde se propagaba el impulso nervioso. Dado que esta estructura formaba un retículo, a esta iniciativa se la denominó como *Teoría Reticular*. Esta teoría fue prontamente aceptada por los sabios de la época, incapaces de demostrar ninguna otra conformación microscópica.

Pero vamos a dejar esta historia aquí, pues quiero comentar otra de las razones por la que estudiar el sistema nervioso era altamente complicado.

Prometo retomar luego el asunto en este mismo punto.

El alto porcentaje de grasa que tiene el cerebro, no solo era un inconveniente a la hora de teñirlo para ver sus componentes celulares, sino que también le confería una consistencia babosa, extremadamente blanda, lo que hacía que fuese muy difícil de manejar. Los cortes finos que se requerían para examinar el tejido al microscopio eran prácticamente imposibles de realizar. Además, el tejido se descomponía rápidamente. Se hicieron muchas pruebas intentando subsanar estos inconvenientes. Así, ya en el siglo XVII (en 1666) Marcello Malpighi (1628–1694) interesado en estudiar el cerebro, hizo una serie de cortes de esta estructura y para endurecerlo se le ocurrió hervirlo. Acto seguido echó tinta sobre las superficies cortada, intentando discernir sus estructuras constituyentes. El método hacía precipitar las proteínas de la pieza y no resultó ser muy afortunado, por lo que en poco tiempo se probó otro método para endurecer las piezas. Robert Boyle (1627–1691) sugiere que se utilice el alcohol para fijar el tejido y da buen resultado. Pero tiene un

inconveniente, pues el alcohol extrae el agua, lo que hace que la pieza pierda volumen y sus estructuras internas se distorsionen. Y esto era un problema que había que subsanar y que se tardó en hacer. En 1809, Johann Christian Reil (1759–1813) dice que la pérdida de volumen se puede solventar en gran medida, añadiendo al alcohol carbonato potásico o amoníaco. Ludvig Levin Jacobson (1783–1843), cirujano en Copenhague, conoce en París a Louis Nicolas Vauquelin (1763-1829), farmacéutico muy interesado en la química. Fue un personaje interesante, que no solo descubre el primer aminoácido (la asparagina), sino también el berilio y el cromo. Pensaba que el cromo podría prevenir la putrefacción de los tejidos y así se lo comunicó a Jacobson. Este, aprovechando esta confianza, se da cuenta en 1833 que una solución de cromato potásico endurecía las piezas de tejido nervioso sin pérdida de volumen y evitando su putrefacción. Pero de la alta afinidad del ácido crómico por el tejido nervioso se dio cuenta otro danés, Adolph Hannover (1814-1894) que trabajaba en la universidad de Berlín con el profesor Johannes Petrus Müller (1801-1858). Trabajando con el sistema nervioso, estaba interesado en encontrar un agente que pudiese preservar, no sólo el aspecto exterior del órgano, sino también su estructura interna. Y después de muchas pruebas, dio la razón a Jacobson porque el único fluido que encontró capaz de satisfacer sus requerimientos fue el ácido crómico. En efecto, no solo protegía de la putrefacción, sino que preservaba la estructura externa e interna del órgano a estudiar (sistema nervioso) y lo endurecía, no demasiado, permitiendo que fuese cortado en finas rodajas para su observación. Así que, al final de muchas vicisitudes que se extendieron en una fracción dilatada de tiempo, las observaciones de Hannover abrieron la puerta a multitud de estudiosos del sistema nervioso. Las soluciones de dicromato potásico se hicieron realmente populares para fijar y preservar los tejidos, especialmente el tejido nervioso. Una vez que tenemos fijado el tejido a estudiar y podemos cortarlo para su observación microscópica, nos enfrentamos al siguiente problema, la tinción. Las piezas requieren ser impregnadas con colorantes selectivos, es decir, con afinidades por distintos componentes, para poder estudiar al microscopio la estructura interna de los tejidos.

Y aquí retomamos la historia donde la dejamos anteriormente.

Recordemos que se ha introducido el carmín como colorante selectivo del sistema nervioso y que este se ha combinado con otros colorantes, como la hematoxilina, para completar la tinción del tejido. En cualquier caso, las impregnaciones no eran buenas y no se conseguía ver satisfactoriamente la morfología y distribución de los procesos celulares. Es entonces cuando entra en escena Camillo Golgi (1843-1926). Este médico e investigador italiano publica el 2 de agosto de 1873 en la “Gazzette Medica Italiana – Lombardia”, el descubrimiento de un nuevo método de tinción del sistema nervioso con tan buena eficiencia que le va a permitir, como él asegura, “desentrañar el estroma intersticial de la corteza cerebral”. Este método, que parecía impregnar las células nerviosas en su totalidad, se basaba en una impregnación con sales de plata, por las cuales el sistema nervioso tiene amplia afinidad. La verdad es que, aunque Golgi estaba intentando encontrar un método idóneo de tinción del sistema nervioso, el hallazgo de este proceder fue pura serendipia. Como he mencionado anteriormente el dicromato potásico era un buen fijador del tejido nervioso y su uso se había extendido entre los histólogos. Golgi solía fijar sus

piezas en este líquido, antes de someterlas a distintos procedimientos tintoriales. Aquella tarde estaba cansado y con ganas de irse a casa. Antes de hacerlo sacó algunas piezas que había estado fijando y que ya habían adquirido ese color anaranjado fuerte, característico del dicromato potásico. Al realizar esta operación, una de las piezas se le cayó en un frasco que tenía abierto sobre la mesa de trabajo y que contenía una solución bastante diluida de nitrato de plata. No se dio cuenta de lo que acababa de ocurrir y se marchó tranquilamente a casa. A la mañana siguiente, poniendo orden en su mesa, se dio cuenta que había algo en ese frasco con la solución de nitrato de plata, que no reconocía. Era una pieza pequeña con un desagradable color negro, ¿sería una de sus piezas de tejido nervioso? Bueno, hasta aquí, la serendipia. Algún investigador, habiendo visto semejante desastre de pieza, la hubiese tirado inmediatamente y continuado recogiendo la mesa. Pero a Golgi le picó la curiosidad, ¿qué habría pasado? Estaba claro que el dicromato potásico, que había penetrado en todos los intersticios de la pieza, había entrado en contacto con el nitrato de plata y había reaccionado, precipitando cromato de plata, que le confería ese color negro. Golgi, se preguntó qué habría pasado en el interior de la pieza, ¿sería todo negro también?... Solo había una manera de salir de dudas, cortar la pieza y mirarla al microscopio. Y dicho y hecho. Cuando Golgi se asoma al ocular de su microscopio, por un momento contiene la respiración; no da crédito a lo que está viendo. El objetivo de su microscopio está mandando a su retina la imagen de una célula aparentemente impregnada en su totalidad. La célula, de color negro destaca claramente contra un fondo amarillo-anaranjado, mostrando su cuerpo celular, dendritas e incluso su axón. Procesos que se pueden seguir a lo largo de la preparación. En el interior de la pieza, el cromato de plata había precipitado sobre la membrana de algunas células haciéndolas visibles en su totalidad. Lo interesante de la tinción era que solo lo había hecho sobre un porcentaje de células bajo, no más del 10%, por lo que las células impregnadas destacaban aisladas y se podía estudiar perfectamente su morfología. Con el tiempo el método se mejoraría y se podrían visualizar pequeños detalles con perfecta claridad. Si el cromato de plata hubiese precipitado, como era de esperar, sobre todas las células que poblaban esa pieza de tejido, Golgi no habría podido ver nada, pues estaría todo negro, dada la alta densidad celular de la sustancia gris de las piezas nerviosas. Al ver esto, Golgi le puso un nombre a este método de tinción: “la reazione nera” (la reacción negra). La Fig. 3 muestra una foto de corteza cerebral impregnada con el método de Golgi.

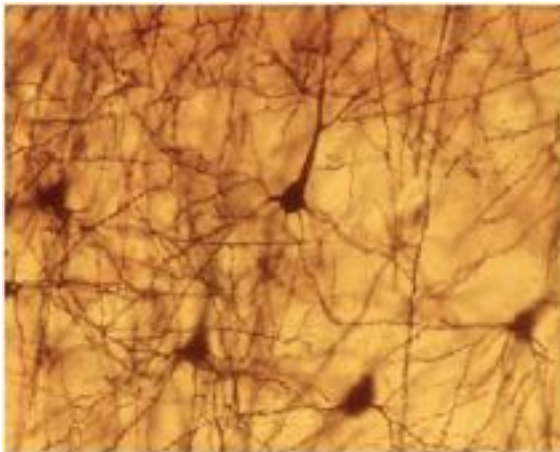


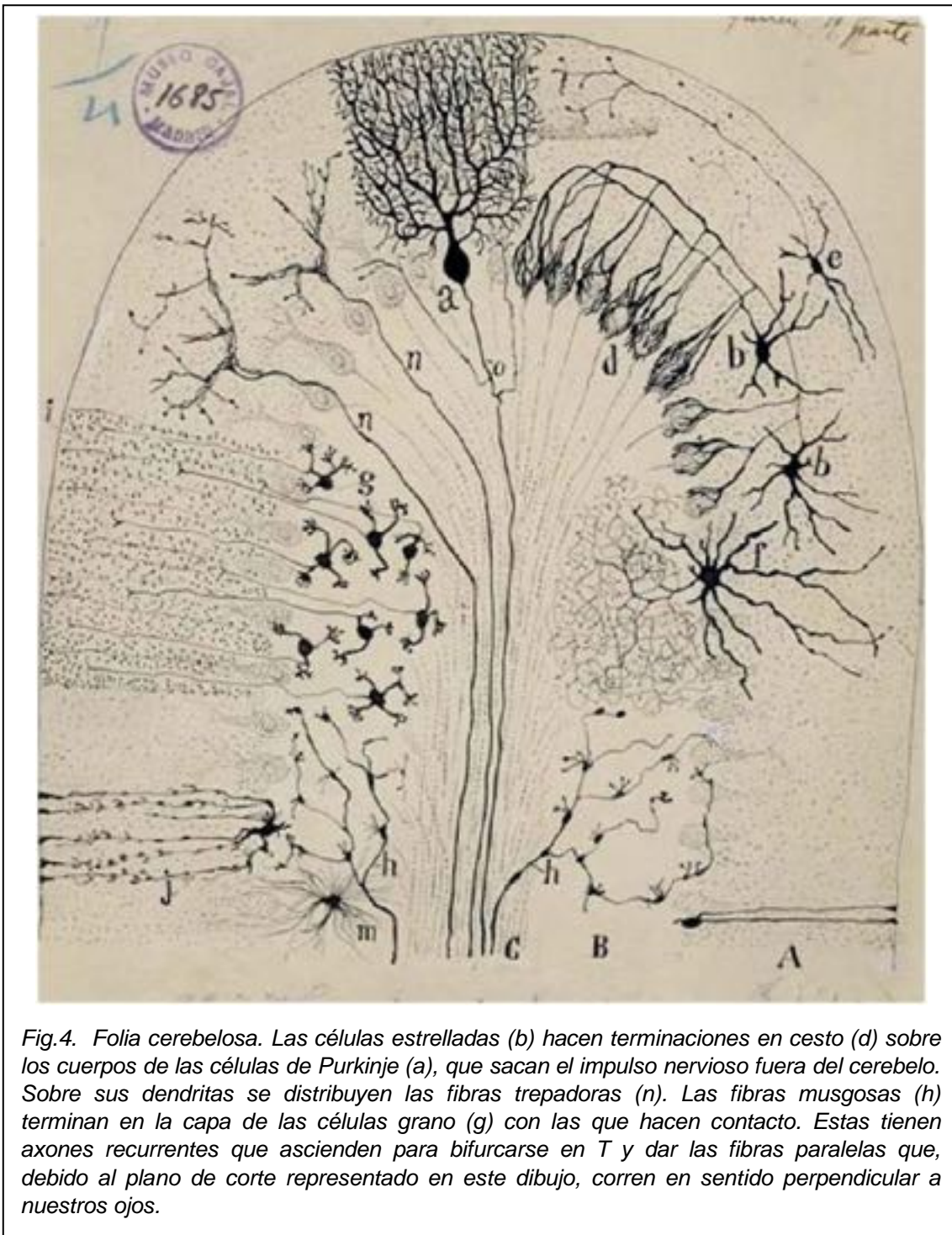
Fig.3. Células piramidales de la corteza cerebral auditiva impregnadas con el método de Golgi. Preparación original de Cajal, conservada en el Legado Cajal (Instituto Cajal-CSIC). Microfotografía tomada por el autor de este capítulo.

Golgi había descubierto un método de impregnación del sistema nervioso sumamente poderoso. Solo tenía que caer en buenas manos. Y digo esto, porque los que hemos trabajado con el método de Golgi, sabemos que no es una panacea. Para empezar, es un método altamente caprichoso, en tanto que impregna lo que quiere, cuando quiere y como quiere. Y hasta el momento, nadie ha podido dar una explicación de ese comportamiento, lo que hubiese implicado poder utilizarlo para impregnar selectivamente tipologías celulares, dependiendo de lo que quisiésemos estudiar. En pocas palabras, para un mismo tamaño de piezas y mismo tiempo de induración en las distintas soluciones, las impregnaciones siempre eran distintas. Unas veces se impregnaban interneuronas y otras veces células de proyección, preferentemente. En otras ocasiones predominaban las fibras y en otras las células y pocos axones. A veces, hay muchos elementos impregnados y otras, muy pocos. No había constancia. A todo esto, hay que lidiar con precipitados, vasos sanguíneos teñidos...es decir, no es fácil de interpretar. Y debió ser lo que le pasó a Golgi, hasta el punto de que se adhirió a las teorías aceptadas en aquella época y usó su método para apoyar vehementemente la Teoría Reticular. Solo introdujo una modificación a la teoría. Si Gerlach había descrito que eran las expansiones protoplásmicas las que se anastomosaban para dar lugar a una red difusa, Golgi, con su nuevo método veía como esos procesos (las dendritas) acababan libremente, por lo que sugirió que los procesos que se anastomosaban eran los cilindroejes o axones. En el mundo científico prácticamente todos eran "reticularistas", es decir, apoyaban la idea de que las células nerviosas se unían para formar un retículo, y estas sugerencias de Golgi fueron aceptadas. Bueno, la verdad es que no todo el mundo. El primero que no pensaba así era Deiters, pero claro, no podía demostrarlo fehacientemente. Pero, de repente, aparecen tres trabajos publicados casi al mismo tiempo, por científicos que no se conocían entre sí y que trabajaban en ciudades distintas. El primero fue Fridtjof Nansen (1861-1930), un joven zoólogo que trabajaba de curador en un museo de Bergen y que, en 1887, escribe un extenso trabajo titulado *The Structure and Combination of the Histological Elements of the Central Nervous System*, donde pone de manifiesto que ha estudiado muchas preparaciones del sistema nervioso impregnadas con tinciones de plata y no ha visto anastomosis entre las células. De caso de existir, estas no se darían como una regla anatómica, sino eventualmente. El segundo artículo proviene de Leipzig y estaba escrito por Wilhelm His (1831-1904), un profesor de anatomía en esta ciudad, de origen suizo. En 1886 reporta que ha estudiado muchos embriones humanos a los que teñía sus tejidos con hematoxilina y que siempre encontraba que, en estadios tempranos del desarrollo de los fetos, el sistema nervioso estaba formado por una masa de células independientes. El tercer artículo procedía de Zúrich y estaba escrito en enero de 1887 por August Forel (1848 -1931), un profesor de psiquiatría que, utilizando la técnica de la "degeneración secundaria", llega a la conclusión de que las células podrán tocarse, pero no se anastomosan entre sí. Piensa esto porque cuando corta unas fibras se produce una degeneración que se limita a las células de las fibras que ha cortado, y no se extiende a células adyacentes. Como vemos, todos estos científicos habían interpretado correctamente sus observaciones, pero dado los métodos de investigación empleados, no tenían nada fácil demostrarlo.

Así están las cosas cuando entra en escena Santiago Ramón y Cajal (1852-1934), un joven catedrático de anatomía en la universidad de Valencia, nacido en Petilla de Aragón, pequeña aldea navarra. Conocido por su apellido materno, Cajal, era muy aficionado a la histología y había intentado estudiar el tejido nervioso en multitud de ocasiones, pero la limitación de los métodos existentes había contenido sus inquietudes científicas. A finales de aquel año de 1887, es llamado a formar parte de un tribunal de oposiciones a cátedra, que se celebran en Madrid. Cuando acude a esta ciudad se encuentra con Luis Simarro (1851-1921), un psiquiatra muy conocido, que acababa de regresar de París, donde había estado trabajando con Ranvier, y que era también muy aficionado a la histología. Cajal consigue entrevistarse con él, y este le enseña unas preparaciones de tejido nervioso impregnadas con el método de plata de Golgi, que le dejan impresionado. Las células nerviosas aparecían teñidas en su totalidad, mostrando cilindro-ejes y prolongaciones protoplásmicas de una manera limpia y clara. Se interesa por los detalles de la metodología llevada a cabo para conseguir tales preparaciones. Piezas de tejido nervioso, no muy grandes (de unos 2 cm) se fijan en una solución de dicromato potásico durante uno o dos meses y luego se transfieren a una solución diluida (0.75%) de nitrato de plata por un par de días. Posteriormente, se deshidrata la pieza, se corta en secciones gordas (150 a 200 micras), se clarifica y se monta en portas de vidrio, cubriendo las secciones con unas gotas de resina dammar (proviene de la *Dammaria orientalis*, árbol de las islas Molucas), sin cubreobjetos. No podía esperar para ensayar el método por sí mismo. Cuando regresa a Valencia se pone manos a la obra. Se da cuenta de lo caprichoso que es el método y decide perfeccionarlo. Una de las innovaciones consiste en incorporar ácido osmico a la solución fijadora de dicromato potásico. Esto hace que se aumente la efectividad y rapidez en el proceso de fijación, de tal modo que las piezas ya no necesitan permanecer en este líquido por dos meses, siendo suficiente una semana. Por este motivo, a esta variedad del método se la denominó "Golgi rápido". Al mismo tiempo, el osmio hace que la impregnación sea más completa y delicada, revelándose más detalles. Intentando aumentar la efectividad, es decir aumentar el número de células impregnadas, realiza doble impregnaciones. Esto consistía en fijar la pieza en dicromato potásico, someterla a la acción del nitrato de plata y en vez de cortarla, volver a empezar el proceso. Los días en las distintas soluciones también los varía dependiendo de la edad del animal utilizado y la zona del sistema nervioso que pretendía estudiar. Esto sería el resultado de numerosas probatinas, donde los frascos y las piezas se acumulaban en las baldas de su laboratorio casero. Con todo esto, Cajal logró domesticar parcialmente el primigenio método de Golgi, aunque no lo logró dominar, como he comentado antes, dado que sigue siendo un misterio su modo de impregnación. Pero Cajal no se contentó con todas estas variaciones e implementó su trabajo poniendo en práctica lo que él denominó como "método ontogénico". ¿Qué es esto? El ocurría de la siguiente manera: Cuando consigo buenas impregnaciones y aparecen muchas células en una preparación de un animal adulto, es como si me adentrase en un bosque muy tupido de árboles, donde me es imposible discernir qué ramas pertenecen a un árbol o a otro, pues están todas muy imbricadas entre sí. Pero, si yo utilizase animales jóvenes e incluso embriones, sería como adentrarme en un bosque que está creciendo y, valga el símil, las células no estarían tan desarrolladas y sería mucho más fácil verlas y estudiarlas. Todas estas modificaciones dieron su fruto, como veremos.

En noviembre obtiene, por concurso de méritos, la cátedra de Histología e Histoquímica Normales y Anatomía Patológica de la Facultad de Medicina de la Universidad de Barcelona, trasladándose a esta ciudad a finales de ese año. Empieza el año de 1888 a trabajar en su nueva cátedra de Barcelona y a estudiar, en su domicilio, el sistema nervioso, de una manera reglada y con el método de Golgi. El órgano elegido para empezar su andadura fue el cerebelo, y los animales, las aves. Bueno, esto era lo más lógico, pues dichos animales los tenía a mano y además eran baratos: pollos y gallinas. Su capacidad de trabajo es grande y el primero de mayo de este año, coincidiendo con su 36 cumpleaños, crea la "*Revista trimestral de Histología normal y patológica*" para publicar sus trabajos, siendo el primero que aparece, un clásico en la literatura neurocientífica, pues describe como los axones de las principales tipologías celulares que pueblan las folias cerebelosas acaban libremente, pudiendo formar diversas estructuras sobre los elementos con los que se ponen en contacto. Así, describe las cestas pericelulares, donde las terminaciones nerviosas envuelven, a modo de cestas, cuerpos celulares de otras neuronas. Describe, asimismo, los dos tipos de fibras aferentes (que entran) a esta estructura: las fibras a las que denomina "musgosas" por la forma en que se distribuyen en la capa de las células grano, y las fibras trepadoras, que literalmente trepan por los cuerpos celulares de las grandes células de Purkinje, para terminar abrazando sus robustas dendritas. Por cierto, sobre estas dendritas encuentra una especie de vello finísimo, con cortos apéndices, a los que denomina "espinas dendríticas". Estas espinas las identificara más tarde sobre las dendritas de las células piramidales de la corteza cerebral. Hoy sabemos que sobre ellas se realizan los contactos sinápticos excitatorios, mientras que los inhibitorios se realizan sobre los segmentos desnudos (sin espinas) de las dendritas o sobre los cuerpos celulares. Por último, identifica a las fibras paralelas, que discurren por la capa más superficial de las folias, como los axones de las pequeñas células grano, que habitan en capas profundas. Con todos estos descubrimientos da una prueba sólida de que las células nerviosas se relacionan unas con otras por contacto y no anastomosándose entre sí. Y aprovecha para establecer una de sus primeras leyes anatómicas, a la que denomina "Ley del contacto pericelular". Esta nueva concepción de cómo se estructura la microanatomía del sistema nervioso, cuestiona las teorías reticularistas vigentes y abren la puerta a la concepción de la "Teoría Neuronal", que establece que la célula nerviosa es una entidad individual anatómica, embriológica y funcional. Y que las neuronas se relacionan entre sí por medio de contactos, es decir por contigüidad y no por continuidad. Posteriormente enunciará otra de sus leyes, la "Ley de la polarización dinámica", que nos explica cómo se transmite el impulso nervioso. En pocas palabras, nos dice que las neuronas están polarizadas, en el sentido que recogen la información de otras neuronas por sus procesos dendríticos, la transportan hacia el cuerpo celular, donde se procesa, y se libera por el axón, que se pone en contacto con otras células para transmitir esa información.

La Fig. 4 muestra un dibujo de Cajal de una folia cerebelosa donde se pueden observar las tipologías celulares existentes en esa estructura y las dos clases de fibras que entran procedentes de núcleos profundos.



Cajal, en el año 1888 quiso publicar sus hallazgos, pero le fue imposible. No dominaba el alemán para redactar un buen trabajo, sus teorías iban en contra de las creencias de aquel tiempo, y nadie le conocía. Conclusión, sus trabajos eran rechazados. Al año siguiente se celebraba en Berlín un congreso internacional de anatómicos y decide ir a enseñar sus preparaciones e intentar convencer a los más reputados sabios de la época que estaban equivocados. En la universidad no le subvencionan el congreso y tiene que pagárselo de su propio bolsillo. Pero merece la pena; conoce al patriarca de la neuroanatomía mundial, Albert Kolliker, al que convencen sus estupendos preparados y sus

explicaciones. Kolliker cree en él y lo da a conocer al mundo científico. Esto hace que Cajal empiece a publicar sus trabajos en las mejores revistas científicas alemanas y francesas y siga cosechando éxitos. Porque Cajal hace grandes descubrimientos anatómicos, pero el genio de Cajal no se encontraba en lo que veía, sino en cómo interpretaba lo que veía. Porque cuando miraba por el ocular de su microscopio, estaba viendo tejido fijado, es decir, muerto; sin embargo, él veía vida donde no existía. El tejido estaba bien conservado, y el dato anatómico hacía funcionar los circuitos cerebrales de Cajal, que se imaginaba cómo funcionaba la estructura que estaba estudiando. Esto permitió el estudio en profundidad del sistema nervioso, de tal manera que podemos afirmar que Cajal es realmente el padre de las neurociencias modernas. Y esto le fue reconocido con creces, pues fue honrado con los premios más prestigiosos que se podían conseguir. Por citar algunos, el premio Moscú (1900), otorgado por el Congreso Internacional de Medicina en París; la prestigiosa medalla de oro de von Helmholtz (1905), que le otorga la Academia Imperial de Ciencias de Berlín; o el premio Nobel en Fisiología o Medicina (1906), otorgado por el Real Instituto Carolino de Estocolmo. En España le proporcionan un laboratorio para que trabaje y cree Escuela, el “Laboratorio de Investigaciones Biológicas”. Posteriormente le construyen otro mayor que por deseo expreso del Rey Alfonso XIII, se llamara “Instituto Cajal”. Le hacen presidente de la Junta para la Ampliación de Estudios; del Instituto Nacional de Higiene, Sueroterapia, Vacunación y Bacteriología “Alfonso XIII”; Senador vitalicio y le quieren hacer Ministro (cargo que rechaza). Cajal, hombre sabio y de gran valía, que logra poner a España en el mapa de la ciencia mundial, es bien reconocido en vida. Tal es así, que se le llega a erigir hasta tres estatuas en vida, hecho totalmente insólito, que solo se había dado en reyes o grandes estadistas, pero en ninguna ocasión en un científico.



Fig.5. Las tres estatuas erigidas en vida a Cajal. (De izquierda a derecha) La primera realizada por Mariano Benlliure en 1922 y emplazada en el Paraninfo de Zaragoza; la segunda realizada por Victorio Macho en 1926, en el Parque del Buen Retiro de Madrid; y la tercera, también en Madrid, en la Facultad de Medicina San Carlos, realizada por Lorenzo Domínguez en 1931.

Cajal ha supuesto un hito para España. Con él se puede hablar que en la ciencia patria y en la mundial, se ha establecido un antes y un después. Por lo tanto, al igual que cuando hablamos del tiempo y queremos datar algún acontecimiento nos referimos como año tal, antes o después de Cristo, cuando hablamos de Neurociencia, nos sería lícito referirnos siempre como antes o después de Cajal.

Bibliografía:

- (1) Alonso, J.R. 2018. Historia del cerebro. Una historia de la humanidad. Editorial Almuzara, Córdoba, España. (ISBN: 9788494778681), 704 pp.
- (2) Alonso, J.R. y De Carlos, J.A. 2018. Cajal. Un grito por la Ciencia. NextDoor Publisher, Pamplona, España. (ISBN: 978-84-947810-9-4), 250 pp.
- (3) De Carlos, J.A. 2001. Los Ramón y Cajal: Una familia aragonesa. Editado por la Diputación General de Aragón, Departamento de Cultura. Zaragoza, España. (ISBN: 84-7753-843-3), 190 pp.
- (4) De Felipe, J. 2014. El jardín de la neurología. Sobre lo bello, el arte y el cerebro. Editores BOE y CSIC, Madrid, España (ISBN: 978-84-00-09897-1), 540 pp.
- (5) Bock, O. 2018. The Morphology of the Nerve Cell: A Nineteenth Century Multi-nation Success Story. Research, 5: 2659 (DOI //dx.doi.org/10.13070/rs.en.5.2659)
- (6) Shepherd G.M. 1991. Foundations of the Neuron Doctrine. Oxford University Press, New York. (ISBN: 0-19-506491-7), 338 pp.

Juan A. de Carlos Segovia

Doctor en Neurobiología.

Investigador Científico, Instituto Cajal – CSIC.

Capítulo 13

La Gravedad: ¡Hágase Newton! Y la luz se hizo... ¿Y Einstein?

*La naturaleza y la luz yacían ocultas en la Noche.
Dijo Dios: "Sea Newton", y todo se hizo luz.
Alexander Pope.*

*Pero no duró mucho, pues el diablo gritó: ¡Ho!,
"Sea Einstein", y se restauró el statu quo.
John Collins.*

El título de este capítulo, como muchos os habréis dado cuenta, hace referencia al elogioso epitafio que el poeta británico Alexander Pope dedicó a Isaac Newton. Menos conocida es la réplica que John Collins le añadió para hacer constar que la maquinaria perfecta de la mecánica newtoniana hubo de ser retocada por Albert Einstein.

De esto es precisamente de lo que os quería hablar, de los primeros pasos de una teoría nueva, la Relatividad General, con los primeros intentos de incorporar la fuerza gravitatoria a la Teoría Especial de la Relatividad. Estamos en los años entre 1905 y 1908. Los argumentos que emplearé son bastante conocidos y sugieren, aunque no describen completamente, fenómenos que solo pueden ser estudiados con la Teoría General. Sin embargo, pese a estar superados hoy en día, dan buenas pistas que apuntan hacia lo que debería describir una teoría completa. Sin llegar a dicha Teoría General, por tanto, me ocuparé de la evolución del llamado *principio de equivalencia*.

Como referencia básica, me he guiado por el capítulo introductorio de [1]. Para las notas históricas, me he basado en [2].

En 1905 Albert Einstein, todavía funcionario de la Oficina de Patentes de Berna, realizó tres contribuciones fundamentales a la Física, que revolucionaron la ciencia del siglo XX.

En el primero de sus trabajos de ese año, Einstein explicó el efecto fotoeléctrico (la emisión de electrones por ciertos materiales al ser irradiados), basándose en la hipótesis de Max Planck, formulada cinco años antes, de que la radiación electromagnética no es un continuo, sino que está formada por paquetes, los cuantos de energía $E = h\nu$ proporcional a la frecuencia de la radiación, ν . La constante de proporcionalidad es la constante h de Planck, que toma un valor muy pequeño, $h = 6.63 \cdot 10^{-34}$ Js, lo que explica que no fueran detectados con anterioridad. Este trabajo le valió el premio Nobel de Física años más tarde. Max Planck había formulado su hipótesis de los cuantos para explicar el espectro de radiación del cuerpo negro. Estos trabajos supusieron la fundación de la mecánica cuántica, la primera de las grandes revoluciones científicas del siglo XX.

En el segundo trabajo, Einstein introdujo una de sus aportaciones más conocidas, la teoría de la relatividad especial, que reconciliaba la mecánica con el electromagnetismo.

Finalmente, Einstein sentó las bases de la mecánica estadística al dar una explicación para el movimiento browniano de las partículas en un fluido.

En este capítulo nos dedicaremos a algunas de las contribuciones a las que estos trabajos dieron lugar. Es decir, nos ocuparemos del trabajo de Einstein posterior a 1905, pero anteriores a la publicación de la teoría general de la relatividad, que permitiría incorporar el campo gravitatorio al marco de la teoría de la relatividad.

Sin llegar a describir la teoría de la relatividad general, analizaremos las consecuencias del llamado *principio de equivalencia*, que simplemente expresa el hecho experimental de que todos los cuerpos caen con igual aceleración, independientemente de su composición, algo ya conocido por Galilei a principios del siglo XVII.

Como veremos, la extensión del principio de equivalencia a otras áreas de la física, distintas de la mecánica, permite extraer notables consecuencias relativas al tiempo y a la trayectoria de los rayos de luz, que precisamente fundamentan la necesidad de trascender la mecánica clásica, y que finalmente conducirían a la formulación, por el propio Einstein, de la teoría de la relatividad general.

Aunque los resultados que se van a consignar son solo cualitativamente correctos, permiten dar una idea de las consecuencias de la relatividad general, sin necesidad de profundizar en las complejidades de la teoría.

Dado que los únicos conocimientos que se van a emplear van a ser el principio de equivalencia y las leyes de Newton, para seguir este capítulo solo son precisas las nociones de física que se imparten en cualquier curso elemental de física general.

Para ello, comenzaremos con una breve revisión de los antecedentes de la mecánica, comenzando por Aristóteles y siguiendo por Galilei y Newton, para dar cuenta de la evolución histórica, a grandes rasgos, de esta rama de la Física, haciendo hincapié en los aspectos relacionados con el principio de equivalencia. Finalmente analizaremos las consecuencias de la extensión de este principio, realizada por Einstein en los años 1907-1908.

La mecánica aristotélica.

Me vais a permitir que me remonte hasta la Edad Antigua, hasta la Grecia clásica. Nuestro punto de partida será la Física de Aristóteles (Estagira 384 a.C.-Calcis 322 a.C.). ¿Por qué tan atrás?, ¿no íbamos a hablar de Einstein?, os preguntaréis. La respuesta es sencilla: porque la influencia de este filósofo se prolongó a lo largo de nuestra Edad Media gracias a la adaptación que hizo Tomás de Aquino de su filosofía para cimentar una teología cristiana, la escolástica. Hasta entonces había primado la influencia de la filosofía de Platón, adaptada al cristianismo a través de las enseñanzas de Agustín de Hipona.

Además, Aristóteles fue uno de los primeros autores griegos en escribir un tratado de Física. Esto no quiere decir que los filósofos anteriores no se ocuparan del tema. Al contrario, desde Tales de Mileto todos los pensadores griegos abordaron con mejor o peor fortuna la *filosofía natural*, el estudio de la naturaleza.

Posiblemente la principal contribución de Aristóteles a la ciencia moderna sea la sistematización de la lógica deductiva, en el sentido de que la ciencia se debe componer de una colección de axiomas, a partir de los cuales, por reglas de razonamiento válidas, como los silogismos, se van deduciendo las proposiciones verdaderas.

En relación con el tema que nos ocupa, la gravitación, para Aristóteles los cuerpos caen *por su propia naturaleza*, buscando su lugar *natural*. Lo mismo que la naturaleza del fuego le hace ascender, lo natural para los cuerpos *graves* sería ubicarse en el centro de la Tierra.

En la citada Física, Aristóteles da pábulo a uno de los errores populares más comunes, que los cuerpos más pesados caen más rápido que los más ligeros. Una idea fuertemente arraigada, que aún hoy no ha sido erradicada de la intuición humana. No pretendamos, por tanto, que lo estuviera en el siglo IV a.C. De hecho, para Aristóteles los tiempos de caída de los cuerpos son inversamente proporcionales a sus pesos. Una afirmación sencilla de rebatir, que sería contestada por Juan Filópono de Alejandría en el siglo VI.

Aparte del movimiento *natural* de los cuerpos, existen otros movimientos, que solo se mantienen por aplicación continuada de una fuerza, ya que el estado *natural* de los cuerpos es el reposo. Por ejemplo, en la filosofía aristotélica un proyectil no sigue una trayectoria parabólica, sino que avanzaría en línea recta hasta perder su impulso, movimiento *antinatural*, para luego descender verticalmente hasta recuperar su posición *natural*, tal como hacen los personajes de los dibujos animados¹⁷.

El principio de equivalencia galileano.

Esta teoría aristotélica del movimiento no fue refutada definitivamente hasta los trabajos de Galileo Galilei (Pisa 1564-Arcetri 1642) a comienzos del siglo XVII. Las aportaciones de Galilei a la mecánica son innumerables. En oposición a Aristóteles, descubrió que los cuerpos, *independientemente de su masa*, caen con aceleración constante y que la distancia recorrida es proporcional al cuadrado del tiempo de caída. Galilei obtuvo estos resultados haciendo rodar bolas por planos inclinados y dejando oscilar péndulos de materiales distintos, pero de igual longitud. La imagen sugerente de Galilei arrojando objetos desde la torre de Pisa es, por tanto, apócrifa.

Aparte de la validación experimental, Galilei aportó un razonamiento, un *Gedankenexperiment*, a favor de su idea: pensemos en dos cuerpos A y B, de masas distintas. Si A es más pesado que B, debería caer más rápido que B, de

¹⁷ https://youtu.be/esR_uxKC27o

acuerdo con la idea preconcebida. Unamos A y B con una cuerda delgada y dejémoslos caer de nuevo. Lo esperable es que B frene la caída de A y por ello el nuevo cuerpo AB debería caer más rápido que B, pero más lento que A. Sin embargo, si seguimos con la idea de que los cuerpos más pesados caen antes, resultará que, como el nuevo cuerpo AB es más pesado aún que los anteriores, debería caer más rápido que A y B por separado. Esta contradicción servía de argumento a Galilei para justificar que la velocidad de caída no podía depender de la masa de los cuerpos.

Este resultado se conoce como la formulación débil del principio de equivalencia o **principio débil de equivalencia**, que podemos expresar afirmando que *dos cuerpos caen, en presencia de la gravedad, al mismo tiempo, independientemente de su masa y de su composición.*

Esta propiedad revela el carácter geométrico *avant la lettre* de la gravedad, ya que, a estos efectos, no importan las propiedades de los cuerpos, sino solo las distancias entre ellos.

Estas y otras contribuciones las recoge Galilei en su obra *De motu* (sobre el movimiento), una colección de ensayos que no llegó a publicar en vida, y en *Discorsi e dimostrazione matematiche intorno à due nuove scienze*, publicada en 1638 en Leyden, Países Bajos, eludiendo así la censura de la Inquisición.

El principio débil de equivalencia fue comprobado con una precisión de un parte en un billón (europeo) por Braginskii y su equipo en los años sesenta, aunque seguramente es más impactante la imagen en vídeo de la caída de una pluma y un martillo en la Luna, realizada por tripulantes de la misión Apolo XV, que podréis localizar fácilmente en la red¹⁸ o el muy gráfico vídeo de la BBC¹⁹.

A Galilei debemos también el **principio de conservación de la cantidad de movimiento**, también conocido como ley de inercia, que afirma que, en ausencias de fuerzas externas, todo cuerpo permanece en reposo o en movimiento rectilíneo uniforme. Es decir, añade el movimiento uniforme a la lista de movimientos *naturales*. Este principio se suele tomar como definición de los **sistemas de referencia inerciales o galileanos**, que son aquellos en los que los cuerpos permanecen en reposo o en movimiento rectilíneo en ausencia de fuerzas.

A pesar de todas sus contribuciones a la mecánica Galilei no llegó a formular una teoría de la gravedad.

Mecánica newtoniana.

La evolución que supusieron las contribuciones de Galilei y Kepler a comienzos del siglo XVII terminó por eclosionar en la figura de Isaac Newton (Woolsthorpe 1643-Londres 1727), quien en su larga vida abordó la práctica totalidad de los

¹⁸ <https://youtu.be/KDp1tiUsZw8>

¹⁹ <https://youtu.be/E43-CfukEgs>

problemas científicos de su tiempo, planteando, a diferencia de sus antecesores, teorías cerradas para los fenómenos naturales, en lugar de descripciones fenomenológicas. Ciertamente, buena parte del pensamiento científico de Newton está presente en otros autores, pero solo él tuvo la genialidad de comprender su alcance y significado. Sin duda sus *Philosophiae naturalis principia mathematica* (1687) constituyen la mayor aportación hasta la fecha a la Física en toda su historia.

En el campo estrictamente matemático, es bien conocida su contribución al cálculo diferencial, compartida en agria polémica con Gottfried von Leibniz. En óptica planteó una bella teoría, la cual, aunque errónea, se anticipó varios siglos a la mecánica cuántica, al sugerir la naturaleza corpuscular de la luz.

En el tema que nos ocupa, nos centraremos en dos trabajos exclusivamente, las tres leyes de la mecánica y la ley de gravitación universal, ambas incluidas en los *Principia*.

La ley de gravitación universal ya había sido postulada para describir el movimiento de los planetas alrededor del Sol por Robert Hooke, quien reclamó su paternidad frente a Newton. Independientemente de ello, el gran acierto de Newton fue precisamente el carácter universal de la ley, pues describe, no solo el movimiento de los planetas, sino el movimiento de cualquier cuerpo sometido a la atracción gravitatoria, como un péndulo o un objeto en caída libre.

La ley de gravitación resume y explica las tres leyes empíricas relativas al movimiento de los planetas de Kepler, quien, como se ha mencionado, no llegó a formular una teoría propia al respecto. Se limitó a expresar leyes fenomenológicas a partir de sus observaciones y las de su predecesor Tycho Brahe, matemático imperial en la corte de Rodolfo II:

1. Los planetas describen órbitas elípticas alrededor del Sol, situado en uno de los focos.
2. El segmento que une el planeta con el Sol barre áreas iguales en tiempos iguales.
3. Para dos planetas cualesquiera el cociente de los cuadrados de sus períodos es igual al cociente de los cubos de los radios medios de sus órbitas.

La genialidad de Newton y de Hooke fue comprender que las dos primeras leyes implicaban una fuerza gravitatoria central F (orientada a lo largo del segmento r que une los centros del Sol y del planeta) y con magnitud proporcional al inverso del cuadrado de la distancia r y a las masas del Sol M y del planeta m ,

$$\mathbf{F} = - (GMm/ r^3) \mathbf{r}$$

si bien Hooke no consiguió demostrar matemáticamente su intuición, logro que corresponde a Newton, quien conocía, no obstante, el trabajo de Hooke.

Al margen de sus éxitos, la ley de gravitación adolecía de un problema filosófico para la época, la acción a distancia. No era concebible que un cuerpo interactuara con otro sin que mediara el contacto entre ellos.

De hecho, René Descartes había formulado con anterioridad otra teoría alternativa de la gravitación, en la cual el universo estaba lleno de un material, el éter, cuyo movimiento en remolinos arrastraba a los cuerpos celestes, los cuales alimentaban a su vez a los remolinos de éter. Esta teoría, modificada por Nicolas Malebranche, estuvo en boga en el continente europeo en competencia con la ley newtoniana hasta que fue finalmente refutada.

En 1735 la Academia de Paris organizó dos expediciones para medir un grado de meridiano de la superficie de la Tierra. Una de las expediciones, comandada por Pierre Louis Moreau de Maupertuis, se encaminó a Laponia, para medirlo en las proximidades del polo, mientras que la segunda, dirigida por Charles Marie de La Condamine, se dirigió a Perú, para medirlo en las proximidades del ecuador. Como consecuencia de las expediciones se estableció que la tierra es oblata (achatada en los polos), en contradicción con la teoría de Descartes, que propugnaba que era prolata (alargada por los polos), y en consonancia con la teoría newtoniana, que recibió el espaldarazo definitivo.

En la expedición al Perú participó el español Jorge Juan y Santacilia (1713-1773), al que recordaréis por los billetes de 10.000 pesetas, cuyas anotaciones fueron recogidas en el libro *Observaciones astronómicas y físicas hechas en los Reinos del Perú*.

Otros grandes éxitos de la teoría newtoniana fueron la predicción de la órbita del cometa Halley en su regreso en 1759, la descripción de las órbitas de planetas y asteroides en el sistema solar con una precisión de una parte en cien millones. Ya en tiempos de Newton se hallaron cometas que seguían órbitas no elípticas, sino parabólicas.

También explica la ley de gravitación universal el fenómeno de las mareas y sirvió para que Urbain Jean Joseph Leverrier y John Couch Adams predijeran la existencia del planeta Neptuno, descubierto en 1846, a partir de las anomalías en la órbita de Urano.

Este último éxito fue tal que cuando se conocieron las discrepancias en la órbita de Neptuno, se postuló la existencia de un noveno planeta, Plutón, que fue descubierto en 1930 por Clyde W. Tombaugh, aunque resultó ser un planeta enano. Del mismo modo, las anomalías en la órbita de Mercurio sugirieron la existencia de un nuevo planeta interior, cuya trayectoria quedaría oculta por el brillo del sol. Sin embargo, en este caso no se halló tal planeta y la explicación correcta tuvo que esperar a la teoría de la relatividad de Albert Einstein.

Así pues, podemos afirmar que la teoría newtoniana murió de éxito, ya que sus logros fueron el motivo de la necesidad de su revisión. Aparte de las discrepancias en la órbita de Mercurio, cuyo perihelio (posición más próxima al sol en su órbita) avanza 42.7" cada año, quedaban otros problemas sin resolver, muchos de ellos de índole filosófica.

Aparte de la *acción a distancia e instantánea* de la fuerza gravitatoria, quedaba por explicar el carácter absoluto de la aceleración en la teoría newtoniana, a diferencia de la velocidad, que solamente es relativa. ¿Respecto

a qué aceleraban los cuerpos? O, dicho de otra manera, ¿qué determinaba los sistemas inerciales? Para Newton la aceleración es relativa al *espacio absoluto*. Otros, como el matemático y filósofo George Berkeley, pensaban que aceleramos respecto a las estrellas fijas lejanas, que determinarían un sistema de referencia inercial. El físico y filósofo Ernst Mach sostenía que era la propia materia del universo la que determinaba la inercia. Aunque este punto de vista motivó a Einstein para formular la teoría de la relatividad, las ideas machianas fueron refutadas por esta.

Principio de equivalencia fuerte.

Saltamos un siglo y nos encontramos con Albert Einstein. Entre 1905 y 1908, fue decisiva su intuición al percatarse de que el principio galileano tenía muchas más implicaciones que las que habían sido ya descubiertas y le otorgó un papel fundamental en su razonamiento. De hecho, las consecuencias del principio de equivalencia son tan profundas que anuncian el final de la mecánica clásica, que paradójicamente había contribuido a fundar.

Del principio de equivalencia se concluye que las leyes de la mecánica son las mismas en un campo gravitatorio uniforme \mathbf{g} que en un sistema con aceleración $-\mathbf{g}$,

$$\mathbf{F} + m\mathbf{g} = m\mathbf{a} \Leftrightarrow \mathbf{F} = m\mathbf{a} - m\mathbf{g}$$

simplemente cambiando de lado el término gravitatorio, convirtiendo la fuerza gravitatoria en aceleración.

Por tanto, podemos eliminar el campo gravitatorio uniforme con una aceleración, por ejemplo, dejándonos caer en caída libre en el propio campo gravitatorio. De hecho, muchos de los experimentos de *ingravedez* se realizan a bordo de aviones o en naves espaciales en órbita con los motores apagados, es decir, en *caída libre* en el campo gravitatorio terrestre. Aunque, por su impulso inicial, no *caigan*, sino que orbiten en trayectorias elípticas. Nuevamente, esto dista de ser obvio popularmente. Solo hace falta echar un vistazo a De la Tierra a la Luna, para encontrarnos con que Julio Verne pensaba que esto solo sucedía en el punto en el que la atracción terrestre y lunar se compensan.

Hasta aquí, no hay nada nuevo que no estuviera presente en el principio de equivalencia de Galilei. La genialidad de Einstein consistió en su generalización.

Si para Galilei no podíamos distinguir entre un campo gravitatorio y una aceleración uniforme con experimentos mecánicos, para Einstein las **leyes de la Física** tampoco deben distinguir entre uno y otra. O, lo que es lo mismo, *las leyes de la Física en un laboratorio en caída libre en un campo gravitatorio son las mismas que en un sistema inercial*. Al menos, a cortas distancias, para que las variaciones del campo en el espacio, las llamadas fuerzas de marea, no sean perceptibles.

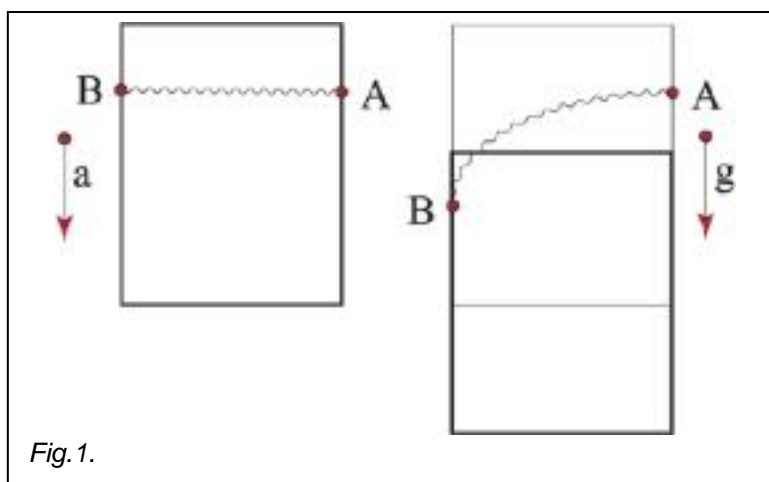
Este enunciado se conoce como **principio de equivalencia fuerte o principio de equivalencia de Einstein** y recoge dos elementos fundamentales: la extensión del principio galileano fuera del ámbito de la mecánica y a campos gravitatorios no uniformes. Veremos a continuación que este principio tiene consecuencias de alcance insospechado.

Curvatura de los rayos de luz.

Consideremos un laboratorio en caída libre (Fig.1), con aceleración, por tanto, igual al campo gravitatorio, $\mathbf{a}=\mathbf{g}$. Desde un extremo A del laboratorio emitimos una señal luminosa.

Como el laboratorio es un sistema inercial de acuerdo con el principio de equivalencia de Einstein, la señal sigue una trayectoria recta dentro del laboratorio en caída libre (parte izquierda de la Fig.1).

Sin embargo, visto desde fuera, es decir, inmersos en el campo gravitatorio \mathbf{g} , observaremos que la trayectoria de la señal se curva por efecto de la gravedad (parte derecha de la Fig.1).



Este resultado tiene sentido si pensamos en la luz como un haz de partículas, los fotones, y tenemos en cuenta que, por la equivalencia entre masa y energía, los fotones han de sentir el campo gravitatorio.

Esta sorprendente predicción del principio de equivalencia fue confirmada después de la primera guerra mundial por una expedición angloalemana a la isla Príncipe, dirigida por el astrónomo británico Arthur Eddington²⁰.

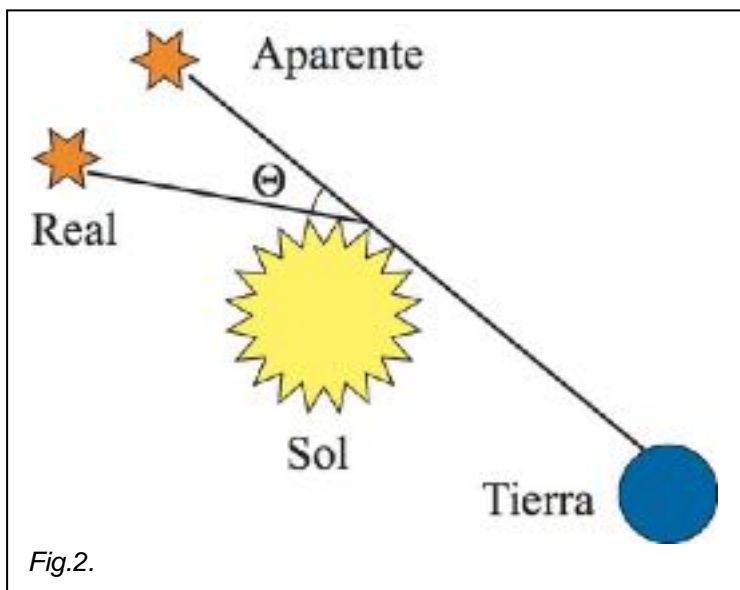
La idea del experimento consistía en que la masa del sol curvaría los rayos de luz (Fig.2) provenientes de estrellas lejanas y por ello se podrían observar estrellas *ocultas* tras el disco solar. El ángulo de desviación esperado era

$$\Theta = (4GM) / (c^2R) = 1.75''$$

siendo G la constante de gravitación universal, M la masa del Sol, c la velocidad de la luz en el vacío y R el radio del Sol.

²⁰ Preguntado Eddington si era cierto que solo había tres personas en el mundo que entendieran la teoría de la relatividad, se quedó un rato dubitativo, preguntándose quién sería el tercero.

En realidad, con el principio de equivalencia solo se predice la mitad de este valor. El resultado completo lo proporciona la teoría de la relatividad general.



Obviamente, la propia luz del Sol impediría la observación de las estrellas, por eso se esperó hasta el eclipse solar de 1919. Intentos anteriores fracasaron por culpa de la guerra, la niebla, la lluvia y la negligencia.

No obstante, la expedición de Eddington tampoco estuvo exenta de emoción. En palabras del propio Eddington:

"La lluvia cesó alrededor del mediodía y hacia la una y media... comenzamos a vislumbrar el sol. Tuvimos que tomar las fotografías a ciegas. No veíamos el eclipse, ya que estábamos demasiado ocupados cambiando las placas, salvo un vistazo para asegurarnos de que había comenzado y otro a la mitad para ver cuánta nube había. Tomamos dieciséis fotografías. Todas eran buenas en cuanto al sol, que mostraba una prominencia destacable; pero la nube había interferido con las imágenes de las estrellas. Unas pocas de las últimas fotografías mostraban unas pocas imágenes que espero que nos proporcionen lo que necesitamos..."

¡De entre las dieciséis tomas, solo una sirvió para confirmar la teoría de Einstein!

Este experimento supuso el espaldarazo definitivo a la teoría de la relatividad general. Hoy en día, repetir el experimento está al alcance de astrónomos aficionados²¹.

Se ha mencionado que Newton había elaborado una teoría, errónea, de la óptica. Sin embargo, su idea de la naturaleza corpuscular de la luz le había llevado a predecir que el sol influiría sobre las trayectorias de los rayos. De hecho, otro físico, Johann von Soldner, halló también el valor del ángulo Θ ya en 1801.

Desplazamiento Doppler gravitatorio.

Otra predicción destacable del principio de equivalencia es que la gravedad no solo afecta a la trayectoria de los rayos de luz, sino también a su frecuencia ν .

²¹ <https://francis.naukas.com/2018/02/05/astronomo-aficionado-eclipse-2017-confirmar-einstein/>

Para justificarlo, volvamos a nuestro laboratorio en un campo gravitatorio g , Esta vez emitimos una señal luminosa desde el techo (Fig.3), la cual, como estamos en un sistema de referencia inercial, tardará en recorrer la altura L del laboratorio en un tiempo $t=L/c$ antes de llegar a un punto A del suelo.

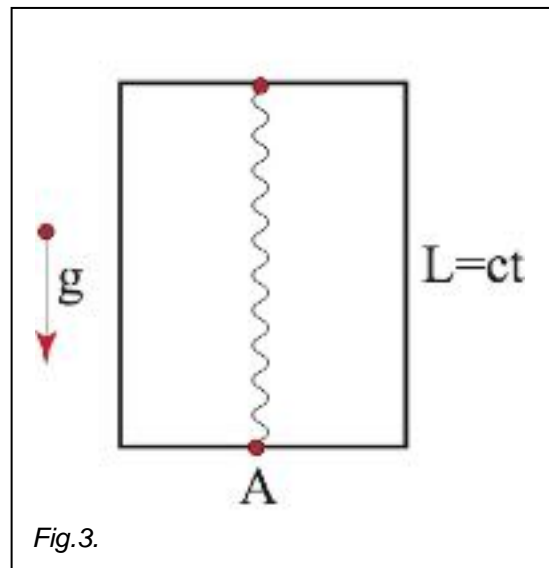
Por la igualdad entre masa y energía, un fotón de energía $E=hc\nu$ pierde energía potencial gravitatoria EgL/c^2 al descender la altura L , que se incorpora a la energía cinética $E=hc\nu_A$ en ese punto inferior. Por tanto, por la conservación de la energía,

$$hc\nu + (hc\nu/c^2)gL = hc\nu_A \Rightarrow (\nu_A - \nu)/\nu = \Delta\nu/\nu = gL/c^2$$

al descender a A en el campo gravitatorio g la señal luminosa experimenta un incremento de frecuencia $\Delta\nu$. Es decir, la señal luminosa sufre un desplazamiento hacia el azul, hacia las frecuencias más energéticas. A la inversa, al remontar un campo gravitatorio, la señal luminosa experimentará un desplazamiento hacia el rojo, al aumentar su energía potencial en detrimento de su energía cinética.

Esta predicción de Einstein no podía verificarse cuando formuló el principio de equivalencia en 1907. Desde los años sesenta, sin embargo, se ha ido comprobando cada vez con mayor precisión hasta llegar a $2 \cdot 10^{-4}$, en un experimento que llevó un reloj basado en un máser de hidrógeno a una altura de diez mil kilómetros.

Este desplazamiento Doppler gravitatorio se corrige de manera rutinaria en los sistemas de posicionamiento global, GPS, basados en relojes que orbitan alrededor de la Tierra.



Dilatación gravitatoria del tiempo.

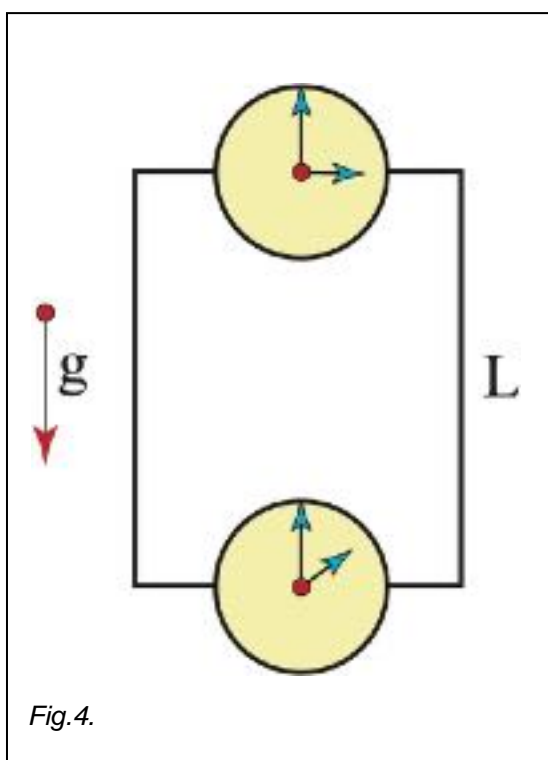
Una consecuencia adicional del desplazamiento Doppler de la frecuencia de una señal en presencia de un campo gravitatorio es la dilatación de los intervalos de tiempo.

Como la frecuencia de la señal es la recíproca de su periodo T , el aumento de la frecuencia al descender en un campo gravitatorio se traduce en una disminución equivalente del periodo T de la señal,

$$T=1/\nu \Rightarrow \Delta T/T = - \Delta\nu/\nu = - gL/c^2$$

Por tanto, el campo gravitatorio dilata los tiempos (Fig.4). Si pensamos, como en el experimento mental anterior, en dos relojes separados una altura L ,

el inferior va más lento, pues mide un intervalo de tiempo menor entre dos crestas de una señal luminosa.



Este fenómeno ha sido comprobado repetidas veces con relojes volando a gran altura y en cohetes. Los relojes en órbita muestran un ligero adelanto respecto a sus homólogos sobre la superficie terrestre.

Es lo que les sucede, por ejemplo, a los protagonistas de la película *Interstellar* (atención, spoiler): al estar sometidos al campo gravitatorio intensísimo de un agujero negro, el tiempo ha transcurrido para ellos más despacio que para sus familiares que se quedaron en la Tierra. O en la recomendable canción '39, de Queen, donde se narra la partida y el regreso de una nave espacial²². Podríamos decir que un campo gravitatorio, o un acelerón intenso, nos mantiene más jóvenes.

Conclusión.

Con este capítulo he pretendido mostrar el gran paso que supuso para la humanidad la extensión que hizo Einstein del principio de equivalencia galileano. Cierto es que no os estoy contando la historia completa, pero a Einstein le proporcionó la intuición necesaria para incorporar el campo gravitatorio a la teoría de la relatividad especial.

Como hemos visto, con la ayuda del principio de equivalencia, Einstein obtuvo resultados que trascendían la mecánica clásica y que motivaron su reformulación.

Sin embargo, el principio de equivalencia no proporciona ninguna teoría consistente, aunque conduce a resultados cualitativamente correctos, por lo cual es de gran ayuda para comprender algunos aspectos de la teoría de la relatividad, sin necesidad de emplear todo su aparato matemático. Este era mi objetivo al redactar este capítulo.

No obstante, para poder comprender los fenómenos a los que apunta el principio de equivalencia, es precisa la teoría general de la relatividad, publicada por Albert Einstein en 1915, tras varios intentos fallidos.

²² Don't you hear my call though you're many years away. Don't you hear me calling you. Write your letters in the sand for the day I take your hand in the land that our grandchildren knew.

Para llegar a ella, Einstein se apoyó, no solo en el principio de equivalencia, sino en el **principio de covariancia**, que afirma que las ecuaciones de la física deben tener la misma forma en cualquier sistema de referencia, inercial o no. Sus ecuaciones son tan conocidas como las de la ya mencionada identidad entre masa y energía,

$$R_{\mu\nu} - (Rg_{\mu\nu})/2 = (8\pi G/c^4)T_{\mu\nu}$$

Donde R es la curvatura del espaciotiempo y T , su contenido de energía y momento.

Ya tendremos ocasión de hablar de ellas, cuando corresponda... De momento, os animo a consular los ficheros de divulgación de la Sociedad Española de Gravitación y Relatividad (SEGRE) en <http://www.segre.es>

Referencias:

[1] Straumann, N., Ed. 2004. *General relativity with applications to astrophysics*. Springer Verlag, Berlin, Heidelberg.

[2] MacTutor history of mathematics archive, University of Saint Andrews.

<http://www-history.mcs.st-and.ac.uk/history/>

Leonardo Fernández Jambrina

Licenciado en Ciencias Físicas y en Ciencias Matemáticas.

Doctor en Ciencias Físicas.

Catedrático de Matemática Aplicada en la Universidad Politécnica de Madrid (UPM).

Expresidente de la Sociedad Española de Gravitación y Relatividad (SEGRE).

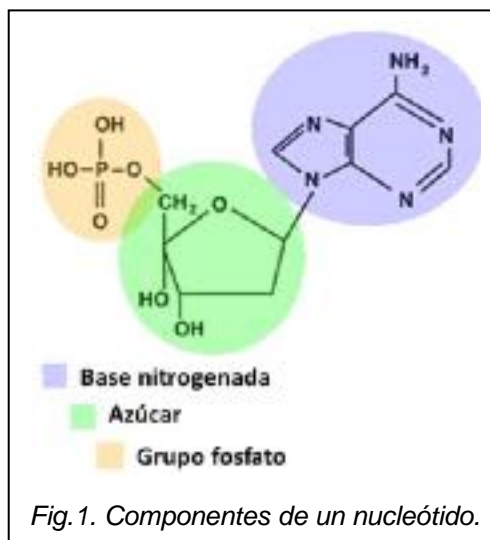
Capítulo 14

Descifrando el genoma humano: la Piedra Rosetta de la biología.

A lo largo de nuestras vidas, siempre hay frases que se nos quedan grabadas en la mente en momentos determinados, que nos hacen reflexionar y que no olvidamos con el paso del tiempo. Un profesor y amigo me dijo una vez “las enfermedades no leen los libros”. Efectivamente, la naturaleza no lee los libros, pero nosotros debemos aprender a leer su lenguaje. Nuestros conocimientos en el campo de la medicina y la biología han ido evolucionando a lo largo de los siglos, rompiendo dogmas con nuevos hallazgos que no han dejado de sorprendernos. El código genético se abrió al ser humano como un nuevo idioma, una fuente inmensa de información que, desde su descubrimiento, no ha dejado de plantear retos tanto científicos como éticos. A día de hoy, los investigadores seguimos intentando entender su lenguaje y seguimos fascinándonos con todo lo que queda por descubrir. Pero para entender mejor estas palabras, empecemos por los cimientos ¿Qué es el ADN?

Descubrimiento del ADN: las primeras piezas del rompecabezas.

A pesar de la creencia popular, Watson y Crick no descubrieron el ADN en la década de los 50. Ellos harían su contribución al conocimiento de esta molécula 80 años más tarde. Fue en 1869 cuando Friedrich Miescher²³, con el fin de extraer e identificar diferentes proteínas a partir de glóbulos blancos procedentes de restos de pus de desechos quirúrgicos, aisló unas moléculas ácidas y ricas en fosfatos presentes en el núcleo de estas células, las cuales poseían características diferentes a cualquier proteína. Miescher denominó a estas moléculas “nucleínas”, término que sería sustituido décadas más tarde por “ácidos nucleicos” [1]. Sin embargo, pasaron más de 50 años antes de que el descubrimiento de Miescher fuera apreciado por la comunidad científica. A pesar de ello, otros científicos, como Phoebus Levene²⁴,



²³ Friedrich Miescher (1844-1895), biólogo y médico suizo. Profesor y rector de la Universidad de Basilea.

²⁴ Phoebus Levene (1869-1940), bioquímico y químico ruso.

continuaron investigando la naturaleza química de estas moléculas. Levene hizo grandes contribuciones en la caracterización de los ácidos nucleicos, siendo el primero en describir en 1919 los tres componentes principales (un grupo fosfato, un azúcar y una base nitrogenada) de las piezas básicas que conforman los ácidos nucleicos: los nucleótidos (Fig. 1) [2].

Como Levene describió, los nucleótidos que conforman el ADN pueden estar compuestos por cuatro bases nitrogenadas diferentes (adenina [A], guanina [G], citosina [C] y timina [T]). Sin embargo, Levene propuso que estos nucleótidos presentaban siempre el mismo patrón de distribución, estando unidos entre ellos siempre en el mismo orden. Como hoy sabemos, esta hipótesis fue desechada, ya que las múltiples combinaciones posibles de estos 4 tipos de nucleótidos sustentan la base del código genético. Fue Erwin Chargaff²⁵, en 1950, quien demostró que los nucleótidos no se distribuían siempre en el mismo orden como había propuesto Levene y que, a pesar de estar presentes en todos los organismos manteniendo ciertas propiedades comunes, la proporción entre ellos variaba en las distintas especies. Además, Chargaff observó que la cantidad de adeninas era siempre muy similar a la cantidad de timinas, y que lo mismo ocurría con las guaninas y citosinas entre sí, estableciendo así la denominada “ley de Chargaff” y sentando una base vital para el posterior trabajo de Watson y Crick (entre otros) respecto a la estructura tridimensional del ADN [3].

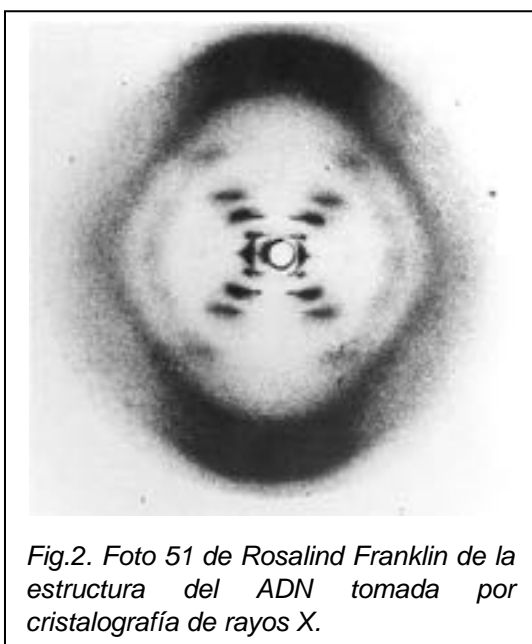


Fig.2. Foto 51 de Rosalind Franklin de la estructura del ADN tomada por cristalografía de rayos X.

Estructura tridimensional del ADN.

En 1952, en el King's college de Londres, Rosalind Franklin²⁶ y Maurice Wilkins²⁷ obtuvieron las primeras imágenes de calidad del ADN por medio de la cristalografía de rayos X²⁸. Wilkins mostró una de las fotos de la molécula de ADN tomadas por Franklin a James Watson²⁹ sin su consentimiento (Fig.2), lo cual lo puso sobre la pista para llevar a cabo la descripción del modelo de doble hélice del ADN. De forma paralela, el químico y bioquímico Linus Pauling³⁰ formuló su propia propuesta con respecto a la estructura

²⁵ Erwin Chargaff (1905-2002), bioquímico y químico austriaco. Profesor en la Universidad de Columbia.

²⁶ Rosalind Franklin (1920-1958), química y cristalógrafa inglesa.

²⁷ Maurice Wilkins (1916-2004), físico originario de Nueva Zelanda. Ganador del Premio Nobel de Fisiología o Medicina en 1962. Trabajó en el Proyecto Manhattan.

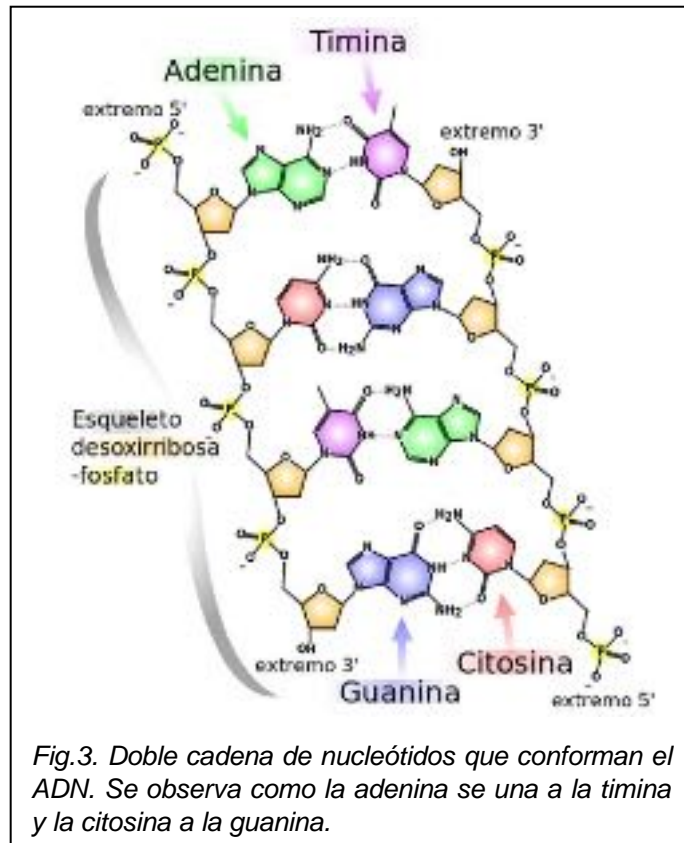
²⁸ Técnica que se usa de manera experimental para el estudio de la estructura de distintos materiales basada en el fenómeno de difracción de los rayos X por sólidos en estado cristalino.

²⁹ James Dewey Watson (1928- hasta la actualidad), biólogo estadounidense. Ganador del premio Nobel de Fisiología o Medicina en 1962.

³⁰ Linus Pauling (1901-1994), químico y bioquímico estadounidense. Ganador del Premio Nobel de Química en 1954 y Premio Nobel de la Paz en 1962.

tridimensional del ADN, la cual, aunque incorrecta, ayudó a la elaboración del modelo actualmente aceptado.

Sustentados en estas bases, James Watson y Francis Crick³¹ emplearon recortes de cartón para representar los componentes químicos individuales de las cuatro bases nitrogenadas, así como de los demás componentes de los nucleótidos, jugando con sus configuraciones atómicas hasta que encontraron una perfecta complementariedad entre la adenina y la timina y entre la citosina y la guanina. Así, estos dos investigadores postularon la estructura en doble hélice de la molécula de ADN, publicando sus hallazgos el 25 de abril de 1953 en la revista *Nature* [4]. De esta manera, se demostró que el ADN está formado por una doble cadena de nucleótidos,



dando lugar a un “efecto espejo” entre ambas cadenas y cumpliendo así la ley de Chargaff ($A=T$ y $C=G$) (Fig. 3). En 1962, Wilkins, Watson y Crick recibieron el Premio Nobel de Fisiología o Medicina por sus estudios sobre la estructura y funcionalidad del ADN, mientras que Franklin falleció en 1958 sin el debido reconocimiento de su trabajo. Sin embargo, a la brillante química y cristalógrafa se le han otorgado numerosos reconocimientos póstumos por sus estudios pioneros, incluyendo la descripción de la estructura del ADN.

Así pues, el descubrimiento del ADN, de sus componentes y de su estructura supuso una larga carrera que marcó un antes y un después en el mundo de la biología. Pero ¿qué información se esconde dentro de las moléculas de ADN y cómo está codificada?

El código genético.

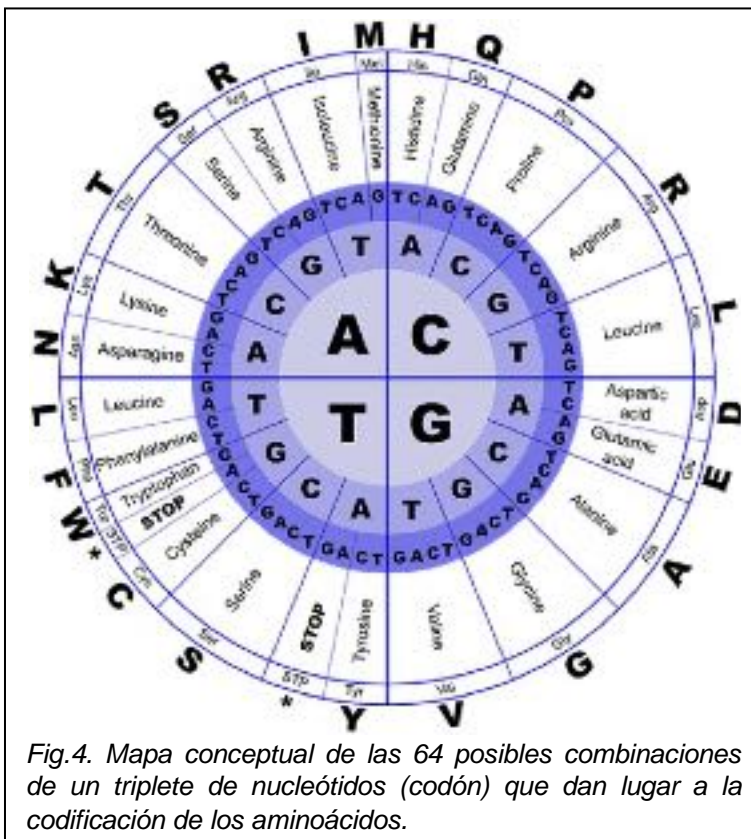
Como ya hemos visto, las diferencias que existen entre los distintos tipos de nucleótidos que conforman el ADN radican en las cuatro bases nitrogenadas. El orden en el que se disponen estas bases nitrogenadas crea un código, un lenguaje cifrado, pero ¿cómo solo 4 letras pueden conformar un idioma tan

³¹ Francis Crick (1916-2004), físico, biólogo molecular y neurocientífico británico. Ganador del premio Nobel de Fisiología o Medicina en 1962.

complejo? Quizás las matemáticas nos puedan arrojar luz sobre esta pregunta. Pensemos en una cadena de 10 nucleótidos, ¿cuántas combinaciones posibles de estas 4 bases nitrogenadas se podrían dar en una cadena de 10 nucleótidos? Inicialmente podría parecer que no demasiadas ¿no? Pues existen 1.048.576 de combinaciones posibles, más de un millón de opciones en solo 10 nucleótidos. Pues bien, dado que se estima que el ADN del genoma humano tiene poco más de 3 mil millones de nucleótidos, las posibilidades que se abren ante nosotros son casi ilimitadas.

De esta manera, el ADN contiene la información necesaria para que cada organismo, cada célula, pueda sintetizar todas las proteínas necesarias en una amplia serie de procesos vitales. Al igual que los ácidos nucleicos, las proteínas también están formadas por piezas básicas, por unidades fundamentales: los aminoácidos. Existen 20 aminoácidos esenciales para la vida, cuyas combinaciones dan lugar a múltiples proteínas con innumerables funciones necesarias para el correcto funcionamiento del organismo. Ahora bien, la pregunta lógica que se nos plantea a continuación es: ¿De qué manera el ADN codifica la información necesaria para sintetizar las proteínas?

Tras la presentación del modelo tridimensional de la molécula de ADN, se comenzó a estudiar en profundidad el proceso de traducción de las proteínas. En este momento, ya se sabía que a partir del ADN podía sintetizarse otro tipo de molécula formada por ácidos nucleicos, el ARN, mediante un proceso llamado transcripción del ADN. A diferencia del ADN, el ARN está formado por una única cadena de ácidos nucleicos y, aunque también está compuesto por 4 bases nitrogenadas, en su estructura la timina es sustituida por una molécula de uracilo. Las moléculas de ARN portan la información necesaria para que se lleve a cabo el proceso de síntesis proteica, denominado traducción del ARN.



es sustituida por una molécula de uracilo. Las moléculas de ARN portan la información necesaria para que se lleve a cabo el proceso de síntesis proteica, denominado traducción del ARN.

En 1953, el físico teórico ruso George Gamow³² quedó fascinado tras leer el trabajo de Watson y Crick y, emocionado, se puso en contacto con ellos inmediatamente para discutir sus ideas sobre tan asombroso descubrimiento. Fue así como, tras una serie de encuentros y con un

³² George Gamow (1904-1968), físico y astrónomo ruso.

desbordante espíritu de colaboración, Watson y Gamow fundaron “el club de la corbata del ARN”. Este club multidisciplinar estaba conformado por 20 miembros (uno por cada aminoácido esencial) y tenía como objetivo averiguar la conexión entre los aminoácidos y el código genético.

Gamow dedujo matemáticamente que, para poder codificar 20 aminoácidos, el número mínimo de nucleótidos necesarios era de 3. Tras la confirmación de esta hipótesis por parte de una serie de experimentos llevados a cabo por Francis Crick y Sydney Brenner³³, entre otros, empezaron a determinarse las correspondencias entre aminoácido y codón (término acuñado por Sydney para el conjunto de 3 nucleótidos que codifica para un aminoácido). Marshall Warren Nirenberg³⁴ y Philip Leder³⁵ fueron capaces de determinar la traducción de 54 codones y, posteriormente, Har Gobind Khorana³⁶ completó el código. De esto se deduce que existen 64 posibles combinaciones de 3 nucleótidos (codones) que darían lugar a la síntesis de los 20 aminoácidos, aunque realmente 3 de estos codones no codifican aminoácidos, si no que indican sitios de parada, es decir, señales para finalizar la traducción de las proteínas (Fig. 4) [5].

La multiplicación de la información genética.

El descubrimiento de la estructura tridimensional del ADN también contribuyó a comprender mejor su mecanismo de replicación, es decir, la manera en la que las células generan copias idénticas de su ADN para que la información genética se transmita de una célula madre a las células hijas. ¿Cómo funciona este proceso? Como ya hemos visto, el ADN está formado por dos cadenas o hebras unidas entre sí a través de las bases nitrogenadas, formando una estructura similar a una cremallera. Esta cremallera se abre por el centro, permitiendo a la maquinaria celular ir sintetizando dos nuevas cadenas a partir de cada una de las mitades de la molécula original, funcionando así cada una de ellas como molde o plantilla para crear una nueva. De esta manera, cada nueva doble hebra tendrá una mitad de la cadena original y una mitad de nueva síntesis (Fig. 5).

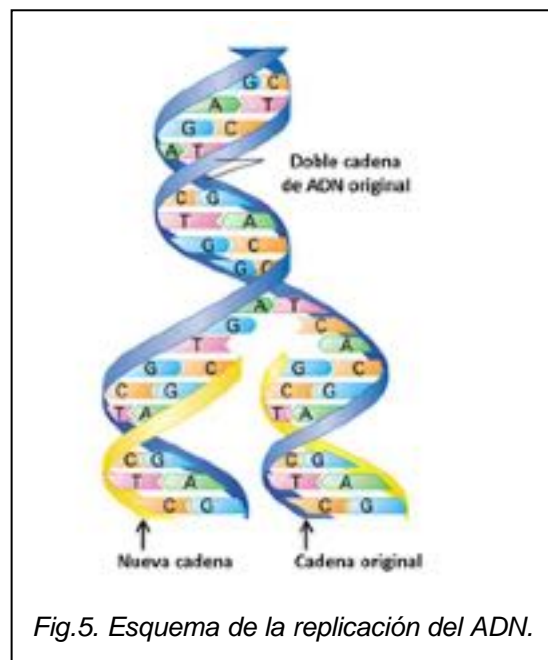


Fig.5. Esquema de la replicación del ADN.

³³ Sydney Brenner (1927-hasta la actualidad), biólogo sudafricano galardonado con el Premio Nobel de Fisiología o Medicina de 2002.

³⁴ Marshall Warren Nirenberg (1927-2010), bioquímico y genetista estadounidense. Ganador del premio Nobel de Fisiología o Medicina en 1968.

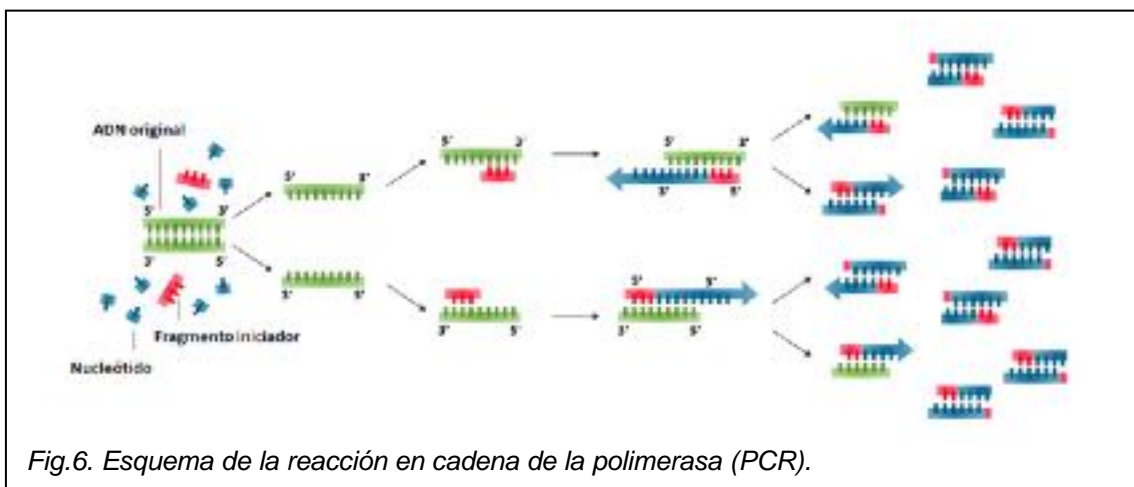
³⁵ Philip Leder (1934-hasta la actualidad), genetista estadounidense. Ganador del premio Nobel de Fisiología o Medicina en 1962.

³⁶ Har Gobind Khorana (1922-2011), biólogo molecular indo-estadounidense de ascendencia panyabí. Premio Nobel de Fisiología o Medicina en 1968.

Sin embargo, la habilidad para “leer” las moléculas de ADN se retrasó en el tiempo. A este procedimiento se le conoce como secuenciación, dado que consiste en conocer la secuencia exacta de los nucleótidos que conforman el ADN, determinando la posición de cada uno de ellos en un orden preciso. Los métodos utilizados para secuenciar los aminoácidos que conforman las proteínas no eran aplicables para esta molécula, dado que el ADN es una estructura mucho más larga y está formada por unidades muy parecidas entre sí, haciendo mucho más difícil la distinción de cada una de ellas de una manera precisa. Por tanto, era necesario el desarrollo de nuevas tácticas que permitieran conocer las secuencias exactas del ADN.

Aunque en la década de los 70 se hicieron algunos avances en las técnicas de secuenciación del ADN, el método desarrollado por Frederick Sanger³⁷ en 1977 supuso un punto de inflexión en esta materia, siendo la metodología más ampliamente usada para secuenciar durante al menos 40 años y obteniendo el Premio Nobel en 1980 gracias a esta innovación. Aunque posteriormente otros métodos han sido descritos y actualmente el primer puesto del pódium está ocupado por técnicas más sofisticadas, el método de Sanger se sigue usando de forma ampliamente generalizada en proyectos de pequeña escala [6].

De esta manera, el conocimiento de los mecanismos de replicación del ADN y el desarrollo de técnicas de secuenciación, entre otros, permitieron que en 1983 Kary Mullis³⁸ desarrollara los primeros esbozos de lo que se convertiría en una de las técnicas más revolucionarias de la biología molecular: La reacción en cadena de la polimerasa o PCR (de sus siglas en inglés “*Polymerase Chain Reaction*”). Esta técnica aprovecha la estructura de doble cadena del ADN para llevar a cabo una amplificación masiva de una zona de interés en el genoma. Empleando dos pequeños fragmentos de ADN que flanqueen la zona que nos interesa amplificar, siendo cada uno de estos fragmentos complementarios a una de las dos cadenas. Además, estos fragmentos funcionan como una señal de inicio de la replicación permitiendo duplicar las cadenas de ADN de forma artificial en el laboratorio.



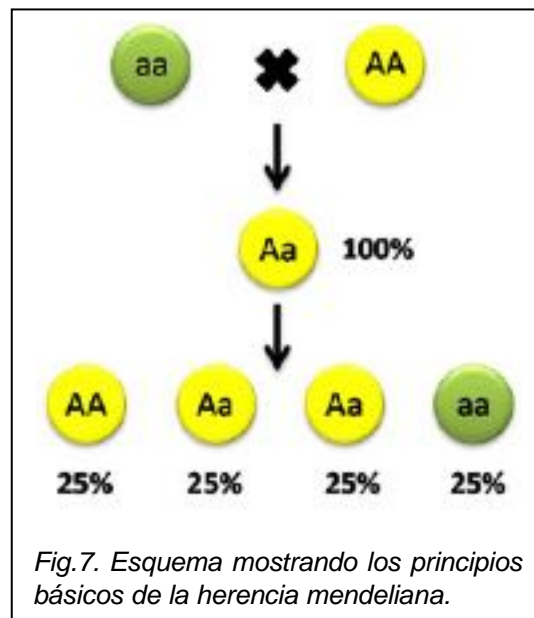
³⁷ Frederick Sanger (1918-2013), un bioquímico británico dos veces laureado con el Premio Nobel de Química (1958 y 1980).

³⁸ Kary Mullis (1944-hasta la actualidad), es un bioquímico, autor y conferenciante estadounidense. Comparte el Premio Nobel de Química de 1993 con Michael Smith.

Ahora bien, ¿Qué ocurre si este procedimiento se lleva cabo de forma repetitiva y cíclica? Que podremos copiar este segmento de ADN de interés una y otra vez, obteniendo muchos millones a partir de una única molécula de ADN en poco tiempo y consiguiendo finalmente una cantidad muy elevada del mismo (Fig. 6). Este sistema abrió una ventana para obtener mucho material genético a partir de una pequeña cantidad inicial ADN, dando lugar a una gigantesca fuente de información de todo el genoma de múltiples especies animales y vegetales, y abriendo innumerables posibilidades futuras [7].

La herencia genética: transfiriendo la información.

Curiosamente, ya se tenía constancia de la herencia del material genético antes de desentrañar la composición y estructura del ADN. Alrededor de 1860 (unos años antes de que Friedrich Miescher aislara la molécula de ADN), Gregor Mendel³⁹ describió que había ciertos “caracteres” que se transmitían a la descendencia en una proporción constante. Observando y cruzando diferentes variedades de plantas de guisantes que diferían en su color y en la rugosidad de su superficie, determinó que algunos de estos caracteres eran predominantes (conocidos como dominantes) sobre los otros (conocidos como recesivos) y que los porcentajes de plantas verdes, amarillas, lisas o rugosas, mantenían siempre la misma relación entre sí. Así, Mendel observó que el color amarillo era dominante sobre el verde y la superficie lisa sobre la rugosa (Fig. 7). De esta manera, el 1865, este monje de origen checo postuló las conocidas como “leyes de Mendel”, explicando cómo estos



“caracteres” se transmitían a la descendencia. Lo que Mendel nunca llegó a saber, es que lo que él denominaba “caracteres” era el reflejo directo de la combinación de los genes aportados por cada uno de los progenitores. Llegados a este punto hay que preguntarse ¿Qué es un gen? Un gen es una secuencia de ADN que es capaz de codificar una molécula que ejerza una función en la célula (ya sea ARN o una proteína). De esta manera, recibimos una copia de cada gen de nuestra madre y otra de nuestro padre, siendo la combinación de ambos lo que determina cómo seremos [8].

Proyecto genoma humano.

Dada la clara importancia médica y social de descifrar el código genético humano, a mitad de los años 80 empezaron a ponerse en marcha diversos

³⁹ Gregor Mendel (1822-1884), monje agustino católico y naturalista.

proyectos con el objetivo de secuenciar y mapear el genoma humano. Con estos antecedentes, en 1990 se fundó el Proyecto Genoma Humano (PGH). Este proyecto internacional, integrado por 20 grupos de investigación de diferentes países, así como por empresas privadas, presentaba dos objetivos principales: determinar la posición de todos nucleótidos que conforman el ADN humano (secuenciación) y localizar la posición de cada gen en cada uno de los 23 pares de cromosomas presentes en las células humanas (mapeo genético). Aunque se estableció un plazo de 15 años para la realización de este proyecto, el 6 de abril del año 2000 se anunció públicamente la terminación del primer borrador del genoma humano, y en 2003, tras varias publicaciones, fue declarado esencialmente completo, dos años antes de la fecha prevista [9].

Desentrañar el conocimiento encerrado en las moléculas de ADN humano ha supuesto grandes ventajas al desarrollo médico y tecnológico en los últimos años, pero también ha planteado diversas consecuencias de carácter ético. Para los investigadores, estos avances supusieron un beneficio tangible, dado que permitieron desarrollar bases de datos en plataformas web, abriendo nuevas vías en sus investigaciones, pudiendo examinar el trabajo de otros científicos en su campo, sobre un determinado gen, sobre su estructura, sobre sus interacciones y funciones, y su implicación en el desarrollo normal del organismo o en la aparición de enfermedades. Este hecho ha permitido un avance vertiginoso en las investigaciones biomédicas a nivel mundial en las dos últimas décadas. Además, gracias al PGH, se ha podido profundizar en el estudio de enfermedades hereditarias, tales como el Alzheimer, la enfermedad de Huntington o el síndrome de Marfan, entre muchas otras. El conocimiento de las bases moleculares de este tipo de enfermedades permite realizar un diagnóstico precoz gracias al cual se pueden tomar medidas preventivas contra la enfermedad o incluso llevar a cabo diagnósticos prenatales para detectar anomalías en el desarrollo del feto antes del nacimiento.

Son muchas las áreas de investigación biomédica que se han beneficiado del PGH. En oncología, poder secuenciar el genoma de un amplio abanico de tumores ha permitido identificar las vías moleculares de muchos tipos de cáncer, dando lugar al desarrollo de tratamientos más personalizados y dirigidos, evitando de esta manera, en la medida de lo posible, la gran cantidad de efectos secundarios que originan los tratamientos basados en quimioterapia convencional. Esta medicina personalizada se ha hecho extensible a muchos campos, permitiendo refinar los tratamientos farmacológicos y desarrollando fármacos que actúen de una manera dirigida según las diversas alteraciones en la secuencia del ADN de determinados genes en los pacientes. En esta misma vía, gracias al PGH, se ha desarrollado lo que se conoce como terapia génica, la cual supone la posibilidad de curar enfermedades de origen genético mediante la modificación directa del ADN del paciente, ya sea introduciendo una copia sana del gen afectado a sus células o directamente modificando la secuencia aberrante de su genoma.

Asimismo, se han realizado grandes avances en el área de la medicina preventiva, es decir, el estudio de un gen o conjunto de genes que pueden predecir la posible aparición de enfermedades que aún no se han desarrollado en el paciente. Sin embargo, aunque en primera instancia esto supone una ventaja para los pacientes, también ha despertado una gran controversia social

desde un punto de vista ético, dado que saber de antemano que determinadas poblaciones o individuos tienen más probabilidad de desarrollar ciertas enfermedades puede dar lugar a la discriminación y rechazo social, así como una desventaja a la hora de conseguir trabajo o un seguro de vida, entre otros. Por ello, ha sido necesario el desarrollo de unos principios éticos que permiten regular estos conocimientos y garantizar la privacidad y dignidad individual, protegiendo la información genética de los pacientes. Además, es necesario potenciar la educación y concienciación de la sociedad en lo que concierne a la información genética, resaltando sus ventajas para el individuo y para la sociedad y clarificando sus derechos de privacidad y decisión individual.

ADN: pasado, presente y futuro.

Como hemos visto a lo largo de este capítulo, desde el descubrimiento del ADN hasta la actualidad han tenido lugar una gran serie de acontecimientos que han revolucionado tanto la manera de pensar de los científicos como la investigación y práctica biomédica. Durante estas décadas, la información que guarda el código genético ha sido minuciosamente estudiada y ha ido revelando un entramado más complejo de lo que inicialmente se creía. Por ejemplo, al principio se pensaba que solo presentaban verdadera relevancia aquellas regiones del ADN que codificaban para la síntesis de proteínas, considerándose la inmensa mayoría del genoma ADN basura. Hoy día se sabe que este pensamiento estaba muy alejado de la realidad y que existen innumerables secuencias de ADN que no codifican proteínas que tienen importantísimas funciones tanto reguladoras como estructurales, así como muchísimos factores que alteran los patrones de expresión de los genes en los distintos tejidos y entre diferentes individuos, dando lugar a múltiples resultados posibles a partir de un mismo gen que no presenta mutaciones. Por ello, no debemos olvidar tener siempre la mente abierta a nuevas posibilidades, nuevos enfoques y descubrimientos. Los avances en la investigación sobre el genoma humano han sido como una lanzadera que nos ha permitido llegar a un nuevo universo antes desconocido, pero aún quedan muchas galaxias por descubrir. El ADN, el genoma, aún no nos ha revelado todos sus secretos, y su amplia complejidad garantiza que aún nos queda mucho por aprender. Ningún buen lector que se precie considera que ya ha leído lo suficiente, y del mismo modo, nosotros seguiremos leyendo y aprendiendo todo lo que el ADN esté dispuesto a revelarnos. El código genético promete ser un “*bestseller*” y nosotros aún no hemos terminado el prólogo.

Referencias:

- [1] Wolf, G. 2003. Friedrich Miescher: The man who discovered DNA. *Chemical Heritage* 21, 10-11, 37–41.
- [2] Levene, P. A. 1919. The structure of yeast nucleic acid. IV. Ammonia hydrolysis. *Journal of Biological Chemistry* 40, 415–424.
- [3] Chargaff, E. 1950. Chemical specificity of nucleic acids and mechanism of their enzymatic degradation. *Experientia* 6, 201–209.
- [4] Watson, J. D., & Crick, F. H. C. 1953. A structure for deoxyribose nucleic acid. *Nature* 171, 737–738.
- [5] Kay L. 2000. Who Wrote the Book of Life?: A History of the Genetic Code. *Stanford University Press*. ISBN 9780804734172.
- [6] Heather, J.M., & Chain, B. 2016. The sequence of sequencers: The history of sequencing DNA. *Genomics*. 107(1): 1–8.
- [7] Templeton, N.S. 1992. The polymerase chain reaction. History, methods, and applications. *DiagnMolPathol*.1(1):58-72.
- [8] Miko, I. 2008. Gregor Mendel and the principles of inheritance. *Nature Education* 1(1):134.
- [9] Chial, H. 2008. DNA sequencing technologies key to the Human Genome Project. *Nature Education* 1(1):219.

Cristian Suárez Cabrera

Licenciado en Veterinaria.

Doctor en Bioquímica, Biología molecular y Biomedicina.

Investigador postdoctoral de la Unidad de Oncología Molecular del Centro de Investigaciones Energéticas, Medioambientales y Tecnológicas (CIEMAT).

Capítulo 15

El Modelo Estándar de la Física de Partículas: un gran paso de la física del siglo XX.

El Modelo Estándar de la Física de Partículas... pero ¿qué narices es eso? Y ¿cómo que fue un gran paso? La llegada a la Luna fue un gran paso, pero esto no lo he oído mencionar en la vida...

Es muy probable que esto sea lo primero que se te ha pasado por la mente al leer el título de este capítulo. Si no ha sido así, ¡enhorabuena! Ya eras un iniciado en física teórica. Si sí ha sido así, tranquilo, nunca es tarde si la dicha es buena ;). Es el momento de aprender sobre uno de los pasos más grandes de la física del siglo XX y de entender de qué va esta fascinante teoría.

Porque, efectivamente, el Modelo Estándar de la Física de Partículas (o Modelo Estándar sin más, a veces incluso SM, del inglés *Standard Model*) es la teoría que clasifica todas las partículas elementales que conocemos y que describe las reglas con las que pueden jugar o comunicarse entre ellas.

Los ladrillos del universo.

Pero espera... ¿qué es una partícula elemental? Esta pregunta ya se la hicieron los griegos hace más de dos mil años. A ellos les fascinaba la naturaleza, todo lo que tenían a su alrededor. Querían entender cómo funcionaba el universo. Y así se plantearon algo tan simple como qué pasa si chocas dos piedras muy fuerte. La respuesta (igual un poco obvia) es que se rompen en trozos más pequeños, simplemente. Pero, ¿y si ahora chocas esos trozos los unos contra los otros muy fuerte? ¿Se romperán en trozos aun más pequeños? ¿Y si esos trozos aun más pequeños los chocas entre ellos muy fuerte...?

Los griegos llegaron a la conclusión de que en algún momento dejarían de poder obtener trozos más pequeños que los que ya tenían, y a estos entes los llamaron átomos, que significa indivisible en griego. Ellos pensaban que lo más pequeño que existía en el universo eran estos átomos, de los cuales había distintos tipos, y que

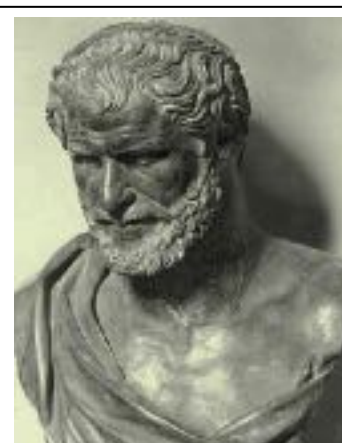
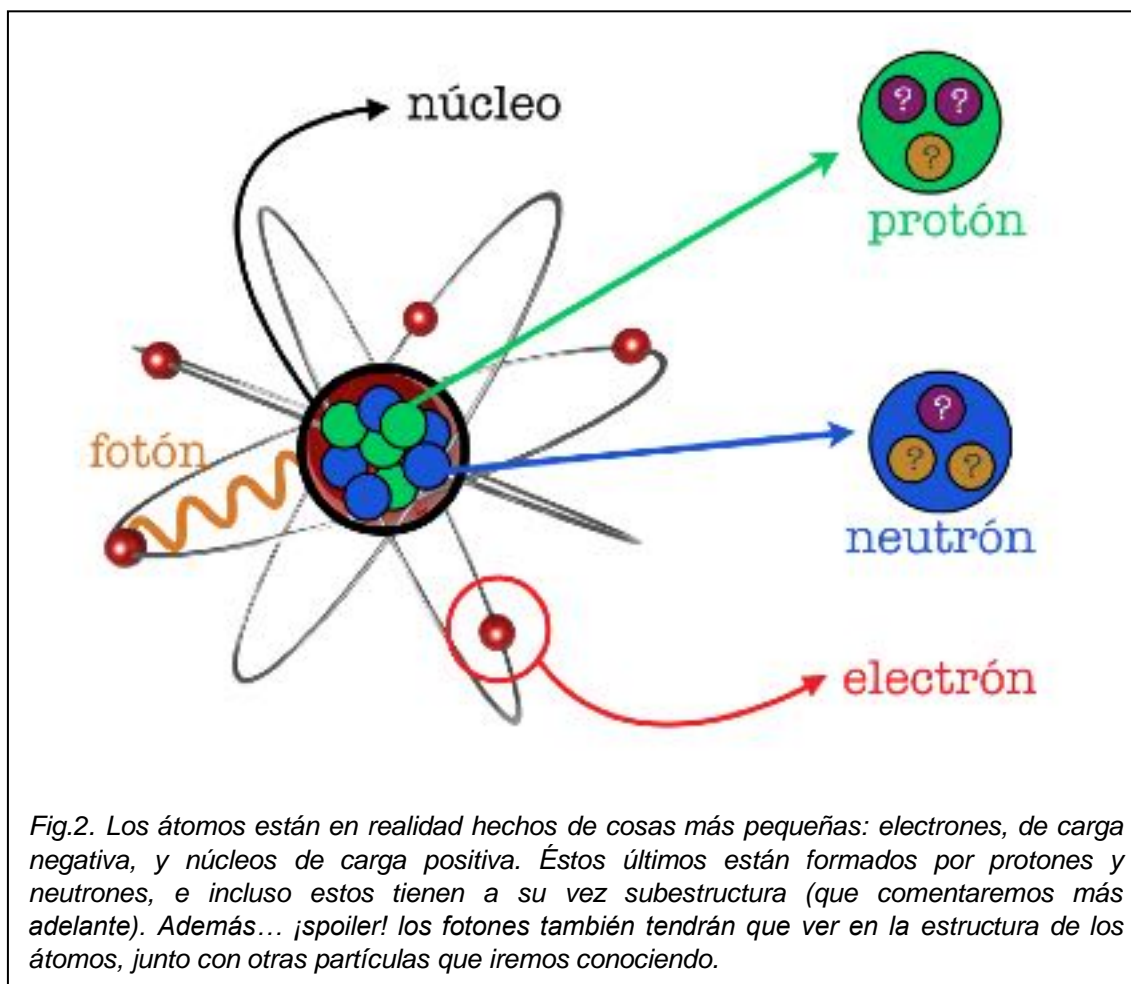


Fig.1. El pensador griego Demócrito fue uno de los primeros que postuló la existencia de algo que no estaba formado de nada más pequeño, a lo que llamó átomo.

combinándolos de distintas maneras podía crearse todo lo que observaban a su alrededor. Es flipante pensar que hace más de dos mil años los malditos genios de los griegos acertaron.

Bueno, no exactamente, pero casi. Al menos los griegos plantearon el concepto de componente o partícula elemental: aquello que carece de subestructura, que no está hecho de nada más pequeño.

Los griegos, sin embargo, se equivocaron un pelín. A finales del siglo XIX y principios del siglo XX se descubrió que los átomos, esos que los químicos y físicos habían pasado siglos estudiando y clasificando en la Tabla Periódica, estaban hechos de cosas más pequeñas. Estas cosas más pequeñas eran los electrones, de carga negativa y los núcleos de carga positiva. Poco después se descubriría, además, que los diminutos núcleos... ¡estaban también hechos de cosas aun más pequeñas! Eran los protones (de carga positiva) y los neutrones (de carga neutra, qué sorpresa, ¿eh?), concentrados en el centro del átomo y en torno a los que orbitaban alegremente los electrones.



Bueno, estupendo, tenemos tres candidatos a ser partículas elementales, los ladrillos que forman todo lo que vemos. Ahora queremos entender cómo se relacionan entre ellos, porque si no menudo aburrimiento,

¿no? Sin embargo, para conseguir comprender cómo interaccionan entre sí estos bichitos escurridizos que son las partículas necesitamos dos ingredientes esenciales nacidos a principios del siglo XX: la mecánica cuántica y la relatividad especial. ¡Tranquilo! Estas palabras suenan a “uff esto es chino, qué complicado”, pero en el fondo solo son dos teorías que nos explican cómo se comportan las cosas muy muy muy pequeñas, como las partículas elementales, y las cosas que van muy muy muy rápido, como la luz, respectivamente. Mmmm... ¿has dicho luz?

La luz son... ¿partículas?

Efectivamente, la luz era el componente que nos faltaba para entender cómo se hablan entre sí los electrones y los núcleos. Resulta que la mecánica cuántica nos dice que existe una dualidad con las ondas y las partículas. Esto quiere decir que, por ejemplo la luz, que en general se entiende como una onda electromagnética, puede ser también, dependiendo de las circunstancias... ¡una partícula! *Mind blown*, lo sé, pero el caso es que este comportamiento dual, estas dos caras que tiene la luz, se han visto experimentalmente. Así que ahora tenemos un candidato más a partícula elemental: la luz, o más concretamente, el fotón, que es la partícula que asociamos con el campo electromagnético.

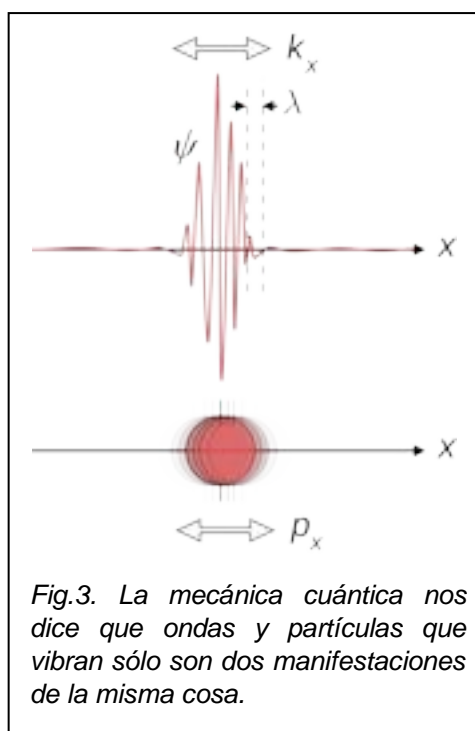


Fig.3. La mecánica cuántica nos dice que ondas y partículas que vibran sólo son dos manifestaciones de la misma cosa.

Además, seguro que alguna vez has oído decir que no hay nada más rápido que la luz. Bueno, pues en general, todas las partículas elementales que conocemos hoy en día, resulta que se mueven a velocidades cercanas o iguales a la de la luz.

Esto hace que para poder entenderlas bien haya que usar la teoría que habla de aquello que se mueve muy muy rápido: la relatividad especial.

Combinando la cuántica y la relatividad surgió la Teoría Cuántica de Campos, que permitía entender bien el comportamiento de las partículas elementales. Concretamente, la Teoría Cuántica de Campos que describía el campo electromagnético (los fotones) y cómo este interacciona con las otras partículas se llamó Electrodinámica Cuántica, o QED del inglés *Quantum Electrodynamics*.

Esta teoría permitió explicar a nivel de partículas elementales por qué las cargas opuestas se atraían y las iguales se repelían. Lo que hacía comportarse a electrones, protones y demás partículas de esa forma era simplemente que se estaban intercambiando un fotón de una manera determinada. Además, durante la construcción teórica de QED, apareció, como por arte de magia, una partícula nueva, exactamente igual al electrón pero de carga opuesta. Esta

partícula, que no era otra que el positrón, era necesaria para que todo tuviera sentido. Poco después se descubrió experimentalmente, demostrándose que QED estaba bien formulada y que existían las antipartículas. Más adelante se descubriría que casi todas las partículas elementales cuentan con su antipartícula, igual a ellas pero con la carga opuesta.

OK, estupendo, ya tenemos nada más y nada menos que ¡cinco partículas elementales! Electrón y positrón, protón, neutrón y fotón. Y además entendemos cómo se relacionan entre sí a través del campo electromagnético. Todo un meritazo, oiga.

Sin embargo, a la vista de los acontecimientos, aun quedan dos preguntas por hacer.

La primera es ¿por qué hay algunas partículas que actúan como ladrillos y otras que actúan como cemento? ¿Hay alguna diferencia entre ellas? Si te paras a pensarlo, de las cinco que ya tenemos las cuatro primeras se observan en la materia (*ladrillos*), dentro de los átomos, pero la quinta, el fotón, se asocia con el campo electromagnético, con una fuerza (*cemento*).

Existe una diferencia entre estas partículas que se llama espín. El espín es una etiqueta intrínseca, una característica, que tienen las partículas y que nos permite clasificarlas. Las que tienen espín semientero, (o sea $1/2, 3/2, \dots$) se llaman fermiones y las que tienen espín entero ($0, 1, 2, \dots$) se llaman bosones. Los fermiones siguen el Principio de Exclusión de Pauli, que dice que no puede haber dos partículas iguales ocupando un mismo sitio. Los bosones, sin embargo, no se rigen por dicho principio y podemos apilonarlos tanto como queramos.

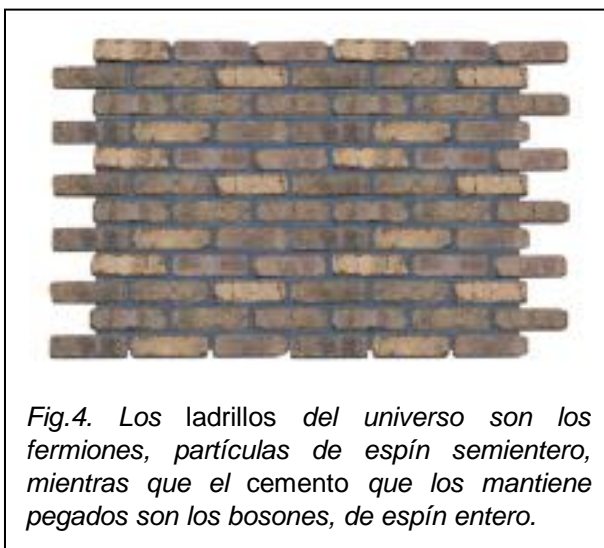


Fig.4. Los ladrillos del universo son los fermiones, partículas de espín semientero, mientras que el cemento que los mantiene pegados son los bosones, de espín entero.

Seguro que ya te has imaginado que los fermiones tienen que ser entonces los ladrillos y los bosones el cemento. ¿Por qué? Pues porque no puedes poner un ladrillo donde ya hay otro, pero si quieres que dos ladrillos queden más pegados siempre puedes echar más cemento. De esta forma se entiende que los fermiones sean las partículas que formen la materia mientras que los bosones sean los mediadores de las fuerzas que la mantienen unida.

¿Por qué no nos desintegramos?

Y precisamente con esto llegamos a nuestra segunda pregunta. Hemos dicho que entendemos bien, gracias a QED, que las partículas de carga opuesta se atraen y las de carga igual se repelen. Esto está muy bien porque con ello

podemos explicar que los electrones negativos orbiten alrededor del núcleo positivo interaccionando con él a través de fotones. Ahora bien, el núcleo está formado por protones y neutrones muy juntitos. Los neutrones no tienen carga, pero los protones son todos positivos... ¿por qué demonios no se repelen?

Para contestar a esta pregunta llegaría la fuerza nuclear fuerte al rescate. Y sí, se llama fuerte simplemente porque para vencer a la repulsión electromagnética entre los protones del núcleo hace falta que la fuerza que los mantiene unidos sea más fuerte. Además, esta fuerza tenía que afectar solo a los componentes del núcleo, y no a los electrones que lo orbitaban.

Tenemos una fuerza nueva, así que, igual que pasaba en QED, deberíamos encontrar partículas asociadas a ella que la transporten, como sucedía con el fotón y el electromagnetismo. Pero estas partículas no aparecían.

Los físicos de mediados de siglo se pusieron a buscarlas como locos... y las encontraron como locos. En pocos años tenían lo que se llamó el Zoo de partículas. Unas de estas partículas eran los piones, que inicialmente se postularon como los mediadores elementales de la fuerza nuclear fuerte. Pero después llegaron un montón de partículas más que no había por dónde coger.

Al fin se obtuvo una respuesta al puzle del Zoo. Los teóricos consiguieron clasificar todas estas nuevas partículas utilizando solo tres componentes distintos, a los que llamaron quarks. Los había de tres "sabores" (el quark arriba o *up* en inglés, el abajo, o *down* y el extraño o *strange*) y cada uno de ellos podía ser de tres "colores". Sí, los físicos son unos graciosillos.

Con estos quarks (que eran fermiones, o sea, ladrillos) se explicaba cómo estaban formadas todas las partículas que se habían encontrado hasta entonces, llamadas hadrones. Incluso se pudo explicar que los protones y neutrones estaban formados también por estos graciosos y coloridos quarks: el protón por dos *up* y un *down* y el neutrón por dos *down* y un *up*. Pero había dos cosas que no terminaban de encajar.

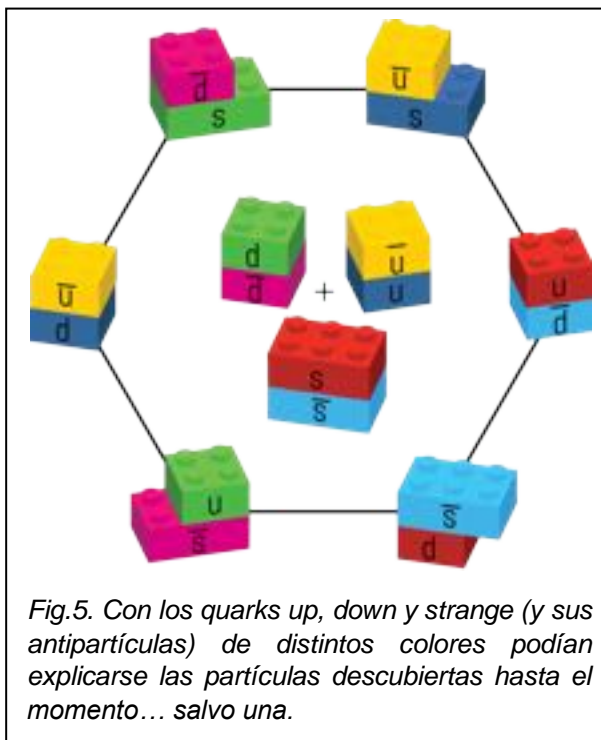


Fig.5. Con los quarks up, down y strange (y sus antipartículas) de distintos colores podían explicarse las partículas descubiertas hasta el momento... salvo una.

La primera era que aun no habían encontrado a los mediadores de la interacción fuerte, que era lo que buscaban en un principio. La segunda es que una de las partículas que habían encontrado era imposible de clasificar utilizando los quarks. Se trataba del muón, una copia idéntica del electrón pero

más gordita (más masiva). ¿Quién demonios había encargado al muón? Además, esa idea de los quarks estaba muy bien, pero si realmente estaban ahí, dentro de los hadrones, ¿por qué nunca se habían visto en los experimentos?

La respuesta a esta última pregunta llegaría algunos años después, con la confirmación empírica de la existencia de los quarks. Si se bombardea un protón con electrones de mucha mucha energía, estos son capaces de “romper” el protón en sus componentes y darnos información sobre su composición. Así se consiguió comprobar que el protón estaba hecho por cosas más pequeñas, lo que cuadraba perfectamente con la explicación de los quarks constituyentes.

Pero, si te has dado cuenta, ¡los malditos mediadores de la interacción fuerte aun no han aparecido! Al final de muchas vueltas los físicos consiguieron entender que estos bosones mediadores, a los que llamaron gluones, estaban dentro de los hadrones al igual que los quarks, y eran muy complicados de ver. La fuerza fuerte era tan fuerte que mantenía los quarks y los gluones confinados dentro de los hadrones, pero estar, estaban.

Todos estos ingredientes formaron, al fin, la teoría cuántica de campos que explicaba por qué los átomos no se iban deshaciendo por ahí debido a la repulsión electromagnética entre los protones del núcleo. La llamaron Cromodinámica cuántica (o QCD del inglés Quantum Chromodynamics), en honor a los colorines de los que habían pintado los quarks para que todo tuviera sentido.

La radiactividad... ¿también son partículas?

Aun así... un momento... sí, lo de que no nos vayamos desintegrando por la vida está muy bien, pero... ¡hay átomos inestables que sí se desintegran! Si no, ¿cómo explicas la radiactividad?

Madre mía, ¿tenemos que meter otra fuerza más para explicar esto? ¡No vamos a acabar nunca por Dios! Tranquilo, ya queda poco, solo nos falta una fuerza, la fuerza nuclear débil, responsable de las desintegraciones nucleares. De nuevo los físicos muy originales con los nombres, le pusieron el nombre de débil ya que esta fuerza tenía que ser más débil que la fuerte para asegurar que mayoritariamente los átomos fueran estables.

Entonces, de acuerdo, vamos a estudiar la radiactividad, pero desde lejos, no nos vaya a pasar lo que a Marie Curie. Existe un tipo de radiación, llamada radiación beta (β), que aparece en algunas desintegraciones atómicas. Resulta que esta radiación no es más que un positrón emitido por un núcleo en el que un protón se transforma en un neutrón.

Al medir las energías del núcleo antes de desintegrarse y después (teniendo en cuenta la del positrón, por supuesto), los físicos vieron que... ¡no se conservaba la energía! Ojo, que eso sí que era un buen marrón, porque la

mítica frase de “la energía ni se crea ni se destruye, solo se transforma” se iba al garete.

Por suerte fueron lo suficientemente listos como para decir, “no, la energía tiene que conservarse, y si parece que no lo hace es porque la que nos falta se la ha llevado algo que es invisible”. ¿*What?* ¿Estás loco? ¿Cómo que algo invisible?

Pues efectivamente, resultó que el positrón que se emitía durante esta conversión de protón a neutrón no se emitía solo. Le acompañaba una partícula extremadamente esquiva, sin carga eléctrica y que apenas interaccionaba con nada. Se la llamó neutrino, y era tan difícil de detectar que parecía invisible. Y la muy condenada era la que se había llevado la energía que faltaba... menuda caradura.

Todo este proceso de la radiación beta tenía que ocurrir a través de la interacción débil, que de nuevo debía llevar asociada alguna partícula mediadora. Pero antes de encontrarla se entendió una característica muy peculiar de la fuerza débil. Resulta que esta interacción funciona como un servicio de citas online, pero en vez de emparejar gente, empareja partículas. De esta forma, el electrón y el tipo de neutrino encontrado en las desintegraciones beta, que se llamó neutrino electrónico por aparecer siempre asociado con un electrón, debían ir juntitos. Los quarks *up* y *down* también. Y el muón y su recién encontrado neutrino muónico también.

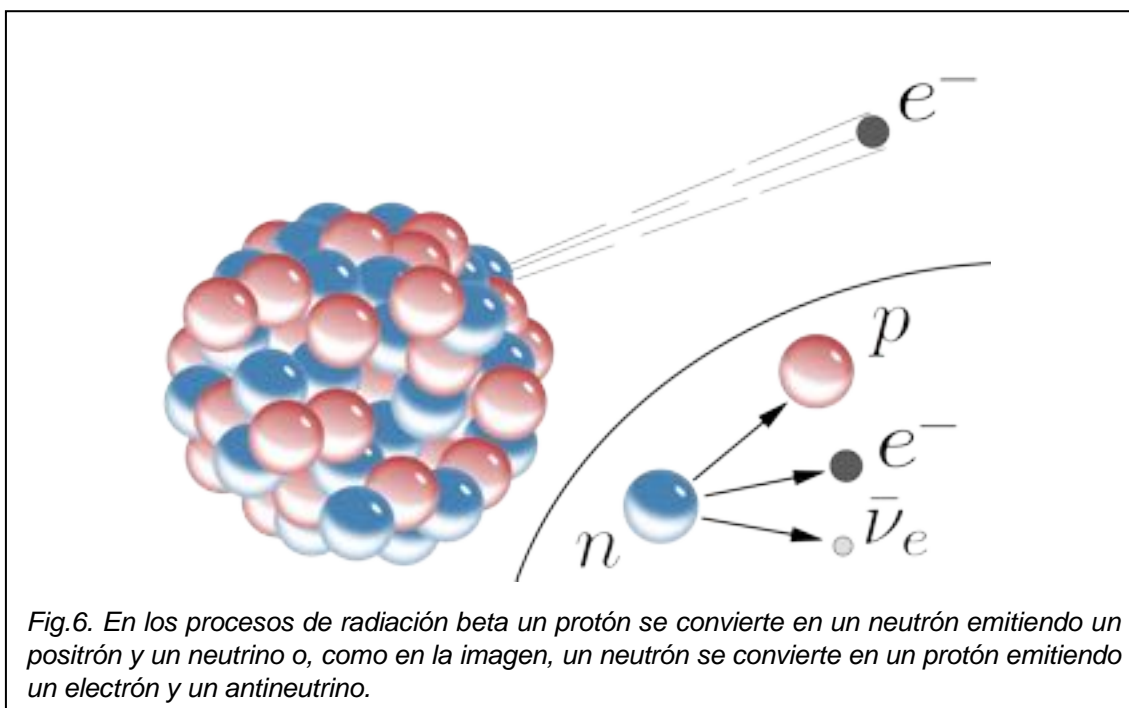


Fig.6. En los procesos de radiación beta un protón se convierte en un neutrón emitiendo un positrón y un neutrino o, como en la imagen, un neutrón se convierte en un protón emitiendo un electrón y un antineutrino.

Pero... se nos ha quedado el *strange* solito... y vale que el pobre es un poco raro, pero no se merece tanta soledad. Necesita un compañero según la fuerza débil. Pues vamos a buscarlo a ver si lo encontramos. Efectivamente, se encontró experimentalmente el compañero del *strange*, y se le llamó quark *charm* (encanto en inglés). Ahora sí, todos bien emparejaditos.

Y además de estar emparejados... qué curioso, parece que hay un patrón, un esquema, una estructura. Si agrupo el *up* y el *down* con el electrón y el neutrino electrónico puedo formar toda la materia que veo y explicar la radiactividad. Y si agrupo el *strange* con el *charm* con el muón y el neutrino muónico tengo una copia idéntica de los cuatro primeros pero más pesada, y como la naturaleza es perezosa a esta copia solo la veo en procesos de muy alta energía. Qué curioso... Y... espera... ¿qué me dices? ¿Que todavía existe una copia más? Exacto, tras conocer los ocho fermiones ya mencionados se encontraron otros cuatro: los quarks *top* (cima) y *bottom* (fondo) y el tau o tauón con su neutrino tauónico. Tío, la naturaleza es muy rara, una copia más y todavía más pesada... ¿para qué?

En fin, muy raro todo. Pero ya tenemos doce fermiones elementales (ladrillos), no está nada mal. También tenemos dos bosones (cemento), el fotón asociado al electromagnetismo y el gluón, asociado a la fuerza fuerte. Y ¿dónde nos habíamos quedado? Ah, ¡sí! En la búsqueda de los mediadores de la débil.

Estos bosones mediadores se habían postulado teóricamente gracias a los procesos de radiación beta. Pero había un problema. Para que todo cuadrara con las observaciones experimentales, dichas partículas —a las que se llamó bosones W, y que vienen en *pack* de 2, uno con carga positiva y el otro con carga negativa— tenían que tener una masa muy grande. Pero la teoría cuántica de campos que explicaba la interacción débil les prohibía tener masa. Y hablando de masas, ¿de dónde sacaban casi todos los fermiones que ya se conocían la suya? Porque masa tenían. ¿De dónde venía?

Y pensarás, me has mentido, bellaca. Me habías dicho que no iba a haber que meter más fuerzas y aquí estamos, con un problema nuevo y sin nada con que explicarlo. Pues no, no te he mentido, no hay que meter más fuerzas, sino algo más sutil y chulo. Hay que romper una simetría.

La maldita partícula (¿O partícula de Dios?).

A los físicos les chiflan las simetrías. Simetría por aquí simetría por allá. Y la verdad es que mola porque con ellas pueden explicarse muchas cosas. De hecho, todas las fuerzas de las que hemos hablado hasta ahora se pueden comprender a través de una simetría asociada a cada una de ellas.

Pues bien, son estas simetrías las que en principio prohíben, y de forma bastante tajante, que las partículas tengan masa. Pero, como ya hemos dicho, muchos fermiones que conocemos tienen masa. Y los mediadores de la interacción débil también. No queda otra pues: la simetría tiene que estar rota.

Sin embargo, si la simetría está rota, tenemos un problema, porque entonces ya no podemos explicar las fuerzas con ella. Maldita sea, y ¿qué hacemos? La respuesta salvadora llegó con el mecanismo de Higgs. Una forma muy particular de romper una simetría sin romperla del todo.

Para entender mejor el concepto por el cual funciona este mecanismo podemos imaginarnos una pelotita en lo alto de una colina. Si la colina es igual por todos sus lados, cuando soltemos la pelotita ésta tenderá a caer hacia los pies de la colina, pero la dirección que escoja para hacerlo será completamente arbitraria. Esto quiere decir que antes de soltar la pelotita existía una simetría que hacía que todas las direcciones posibles fueran iguales. Sin embargo, al soltarla, la pelotita cae en una y solo una de las direcciones, rompiendo la simetría.

De una forma parecida funciona el mecanismo de Higgs. Lo que hace este mecanismo es romper sin romper la simetría electrodébil. Porque, a estas alturas de la película, se había logrado comprender que la fuerza débil era tan romántica que no le gustaba nada estar sola, así que se la emparejó con la electromagnética dando lugar a la fuerza electrodébil. Esta unificación fue algo similar a lo que ocurrió con el electromagnetismo a finales del siglo XIX: dos cosas que parecían distintas, los fenómenos eléctricos y los magnéticos, en realidad eran dos manifestaciones de la misma cosa: el electromagnetismo. Lo mismo le pasaba a las interacciones electromagnética y débil, en realidad eran las dos caras de la interacción electrodébil.

El mecanismo de Higgs funcionaba estupendamente, le daba masa a los fermiones y a los bosones W sin cargarse nada de la bonita teoría que explicaba el comportamiento de todas estas partículas. Perfecto. Y además traía consigo sorpresas en forma de hipotéticas nuevas partículas.

La primera era un compañero de faena de los bosones W, pero esta vez, neutro. Se llamó bosón Z, y se confirmó su existencia experimentalmente a principios de los años 80. Su masa también podía explicarse a través del mecanismo de Higgs, así que todo funcionaba a las mil maravillas.

La segunda era el famoso bosón de Higgs, llamada, erróneamente, la partícula de Dios. En realidad esta partícula había de llamarse *The Goddamn Particle*, la maldita partícula, pero *The God Particle*, la partícula de Dios, quedaba más chachi, así que se quedó con ese nombre. Este bosón era distinto a los demás bosones que ya conocemos. ¿Por qué? Porque el bosón de Higgs no es mediador de ninguna interacción, y, además, tiene espín cero, mientras que el fotón, el gluón, los W y el Z tienen espín uno.



Fig.7. El bosón de Higgs era la última pieza que daría sentido a todo el Modelo Estándar. Pero había que encontrarla experimentalmente y no era tarea fácil. Para ello, entre otras cosas, se construyó la máquina más grande creada por el hombre: el Gran Colisionador de Hadrones.

Era imprescindible que el bosón de Higgs existiera para que el mecanismo de Higgs fuera el que realmente explicaba bien la masa de las partículas elementales conocidas. Sin él, que era una predicción directa de este

mecanismo, no habría confirmación de que éste funcionara, y habría que pensar en otra cosa.

Pasaban los años y el bosón de Higgs no aparecía. Y eso que se hicieron esfuerzos hercúleos para encontrarlo. De hecho, se construyó la mayor máquina jamás creada por el hombre. Un acelerador de partículas de veintisiete kilómetros de circunferencia enterrado bajo la ciudad de Ginebra. O lo que es lo mismo, un microscopio tan potente que podía ver las partículas elementales. Si se iba a encontrar el bosón de Higgs, tendría que ser con esa máquina, el Gran Colisionador de Hadrones o LHC, del inglés *Large Hadron Collider*.

Pero no había manera, y los físicos se desesperaban... Hasta que, por fin, el cuatro de julio de 2012... ¡Se descubrió el bosón de Higgs! ¡¡¡Yupi!!!

De esta manera quedó por fin completo el Modelo Estándar, que recoge las partículas elementales que conocemos a día de hoy y las interacciones entre ellas.

El Modelo Estándar de la Física de Partículas.

Ha sido un duro y largo viaje a través de teorías locas y muchos quebraderos de cabeza. De simetrías que se rompen sin romperse y de fuerzas que son a la vez partículas. Pero al final lo hemos conseguido, hemos construido el Modelo Estándar. Minipunto para nosotros.

Para dejarlo todo claro y tenerlo juntito, hagamos un resumen de esta “Tabla Periódica Moderna”:

Por un lado, tenemos los ladrillos, los fermiones, las partículas de materia que conforman todo aquello que vemos. Son doce, y están organizados en tres familias. Tres copias idénticas cada vez más pesadas.

En la primera familia tenemos dos quarks, el *up* y *down* y dos leptones (que así se llaman los fermiones que no son quarks, es decir, que no sufren la interacción fuerte), el electrón y el neutrino electrónico.

En la segunda, una copia formada por los quarks *charm* y *strange* y por el muón y el neutrino muónico.

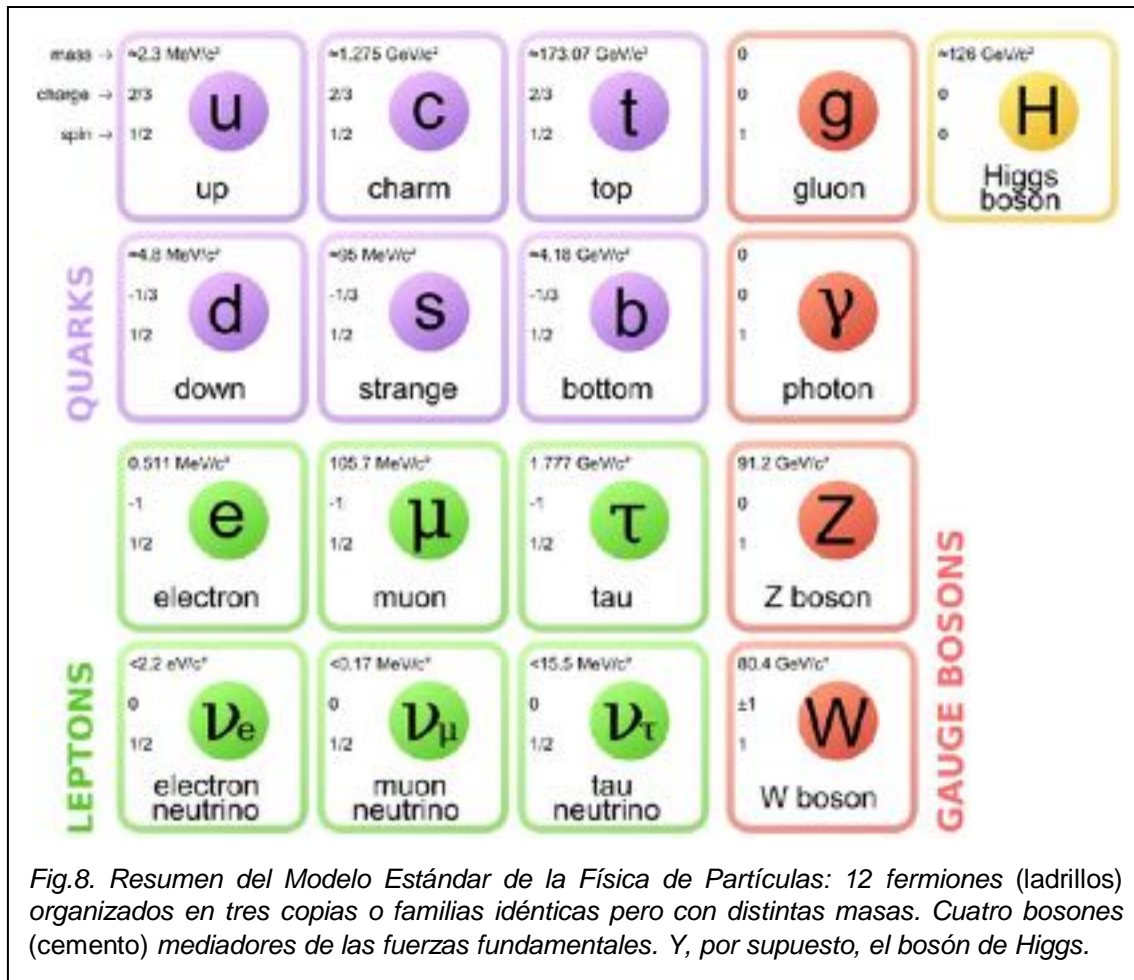
Y finalmente, en la tercera, tenemos a los quarks *top* y *bottom* y al tau y al neutrino tauónico.

Recordemos también que por cada una de estas partículas tenemos que tener también en cuenta su antipartícula, igual a ella pero con las cargas opuestas.

Por otro lado, tenemos a los bosones de espín uno, los mediadores de las fuerzas. Están el fotón, asociado a la fuerza electromagnética, el gluón,

asociado a la fuerza nuclear fuerte y los bosones W y Z asociados a la fuerza nuclear débil.

Además de todos estos, tenemos a nuestro bosón peculiar, el de Higgs, que es el responsable de dar masa a las partículas del modelo.



¿Hay algo más allá?

Y ya está, ya hemos acabado, ya sabemos lo que es el Modelo Estándar. Con él podemos predecir y calcular de forma extraordinariamente precisa cantidades relacionadas con la física de partículas. Y además estas cantidades están increíblemente de acuerdo con los resultados experimentales.

Así que colorín colorado, este cuento se ha acabado. Ya tenemos la teoría que explica perfectamente cómo funciona el mundo de las partículas elementales... ¿O no?

Sorprende, pero no. Afortunadamente, aun nos quedan muchas cosas por descubrir. Es cierto que el Modelo Estándar es una teoría increíblemente precisa, muy bien formulada, y capaz de explicar numerosos procesos en física de partículas, en la física de los bichos más pequeños del universo. Pero aun

hay muchos fenómenos que no terminamos de entender y para los que el Modelo Estándar no tiene respuesta.

Si te has quedado con ganas de saber más, te recomiendo que leas el capítulo de este mismo libro “El más allá del Modelo Estándar”, escrito por Xabier Marcano. En él podrás descubrir cuáles son algunos de esos fenómenos que el Modelo Estándar no es capaz de explicar, y las posibilidades que se barajan para entenderlos.

¿Nuevas interacciones? (Que no, que no te he mentado, esas ya estarían fuera del Modelo Estándar y yo prometí contarte cosas solo acerca de este)
¿Nuevas partículas de materia? ¿Por qué está todo hecho de materia y no de antimateria? ¿Esconde el bosón de Higgs algo más? ¿Qué peculiaridades tienen los neutrinos?

Estas y muchas más preguntas son las que los físicos intentan entender cada día, guiados por la misma curiosidad y ganas de saber que los llevó hasta crear una de las teorías más precisas y bonitas de la historia de la ciencia: El Modelo Estándar de la Física de Partículas.

Claudia García García

Estudiante de Doctorado en Física Teórica.

Universidad Autónoma de Madrid e Instituto de Física Teórica UAM/CSIC.

Capítulo 16

La Inteligencia Artificial: el largo camino desde la IA específica a la IA general.

Imaginemos que tuviéramos una máquina para viajar en el tiempo y transportáramos a Isaac Newton desde finales del siglo XVII hasta la actualidad para situarlo en un lugar que le resultara familiar como por ejemplo la capilla del Trinity College en la Universidad de Cambridge. Una vez allí imaginemos que le mostramos un móvil de última generación y lo conectamos. Sin duda, él que demostró que la luz blanca se descompone en colores al incidir un rayo de sol en un prisma, se sorprendería de que un objeto tan pequeño produjera colores tan vivos en la oscuridad de la capilla. Después hacemos que en el móvil suene una música que seguramente reconocería, por ejemplo una ópera de Händel. A continuación le mostramos en la pantalla su obra “Principia Mathematica” y le hacemos ver cómo usar dos dedos para ampliar el texto. Supongamos también que acto seguido le mostramos como hacer fotos, grabar videos y sonido, hacer cálculos aritméticos con gran velocidad y precisión, contar los pasos que andamos, guiarnos hacia nuestro destino y, por supuesto, hablar con alguien a miles de kilómetros. ¿Sería capaz Newton de dar una mínima explicación de cómo funciona tan maravilloso dispositivo? A pesar de ser una de las mentes más brillantes de la historia, que inventó el cálculo infinitesimal e integral, explicó tanto la óptica como la gravedad y formuló las leyes del movimiento de los cuerpos que revolucionaron la Física, sería incapaz de dar una explicación mínimamente coherente. No podría distinguir ese dispositivo de la magia. Por lo tanto, ¿qué más se imaginaría Newton que ese dispositivo pudiera hacer? ¿Creería que puede funcionar indefinidamente? - recordemos que vivió en una época 100 años anterior a Alessandro Volta, el inventor de la pila eléctrica-, ¿creería que puede transformar plomo en oro? - recordemos que la Química de su época era la Alquimia-. Posiblemente sí, ya que tendemos a no ver los límites a aquello que nos parece mágico. Este es precisamente uno de los problemas que tenemos a la hora de comprender tecnologías muy avanzadas. Ya lo dijo Arthur Clarke en los años 60: “Cualquier tecnología lo suficientemente sofisticada no puede distinguirse de la magia”. En Inteligencia Artificial ocurre lo mismo. Parece que no haya límites en su potencial pero en realidad la IA permanece atascada desde hace más de 50 años en una de las cuestiones más fundamentales: ¿Cómo dotar de conocimientos de sentido común a las máquinas? Es una cuestión crucial si queremos lograr inteligencias artificiales de tipo general indistinguibles de la inteligencia humana. Hasta hoy día los investigadores en IA no vemos ningún indicio que nos lleve a poder afirmar que este problema pueda ser resuelto ni a corto ni a medio plazo. De hecho DARPA, la institución que más invierte en programas de investigación en EEUU, a finales de 2018 anunció que iba a financiar con 2000 millones de dólares un programa de investigación sobre

como dotar de conocimientos de sentido común a las máquinas. La ausencia de conocimientos de sentido común imposibilita que un sistema de IA pueda comprender el lenguaje, pueda comprender lo que percibe mediante sus sensores, pueda manejarse bien en situaciones imprevistas y pueda aprender a partir de la experiencia. Resolver el problema de la adquisición de conocimientos de sentido común sería un gran avance de la inteligencia artificial pues abriría la puerta al desarrollo de inteligencias artificiales de tipo general superando de esta manera las limitaciones actuales de la IA específica, es decir capaz de llevar a cabo únicamente una sola tarea.

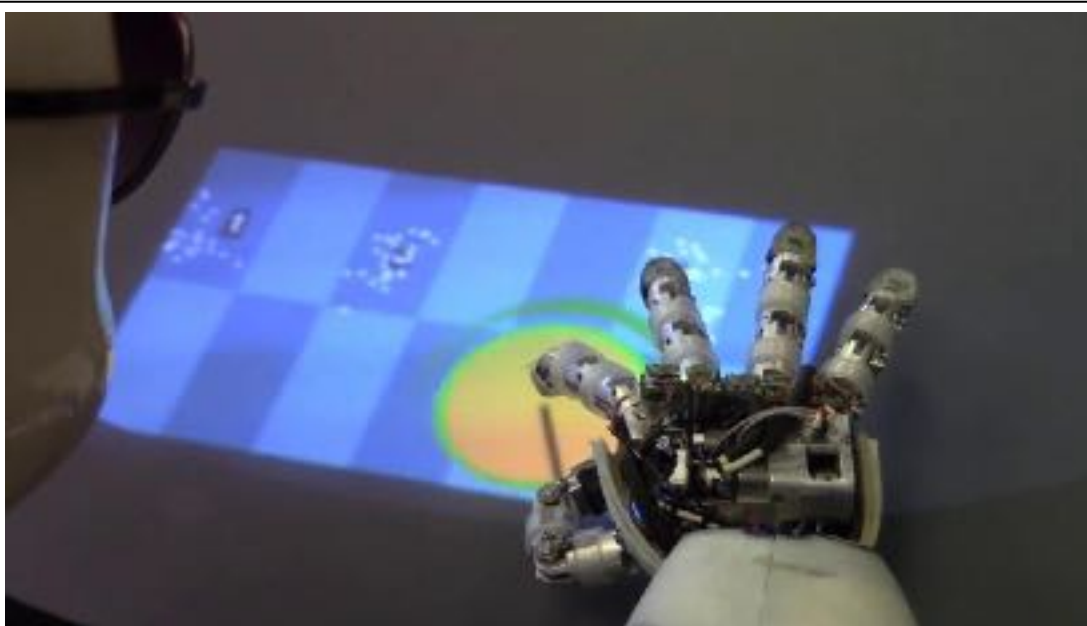


Fig.1. Robot humanoide experimentando con un teclado musical virtual para aprender la posición donde suena cada nota. Crédito IIIA.

¿Cuál es pues la situación real de la IA? La realidad es que lo que tenemos son “inteligencias” sumamente específicas en el sentido que cada una de ellas sabe hacer bien una única tarea. Focalicémonos en una técnica de IA, conocida como “aprendizaje profundo” que ha conseguido espectaculares resultados recientemente como por ejemplo un software llamado Alpha Zero que, jugando contra sí mismo millones de partidas durante horas, aprendió a jugar a Go a unos niveles nunca antes alcanzados superando con creces a los mejores jugadores humanos. Pues bien, estos sistemas de aprendizaje profundo son sumamente limitados ya que únicamente son capaces de aprender a clasificar patrones analizando enormes cantidades de datos. No es exagerado afirmar que, de hecho, no aprenden realmente nada; por lo menos en el sentido humano de lo que entendemos por aprender. Es decir que en realidad no saben nada nuevo después de haber sido entrenados para adquirir una competencia. Prueba de ello es lo que se conoce como “olvido catastrófico” que significa que los sistemas de aprendizaje profundo olvidan todo lo aprendido previamente a partir del instante que se les enseña algo nuevo. Por ejemplo, si después de haber “aprendido” a jugar a Go entrenamos a un sistema de aprendizaje profundo a diferenciar entre gatos y perros en base a mostrarle millones de imágenes de ambos, aprenderá perfectamente a

distinguirlos pero será incapaz de volver a jugar a Go. Sería necesario volver a entrenarlo para que de nuevo “aprendiera a jugar a Go” lo cual provocaría que a continuación sería incapaz de distinguir los gatos de los perros. En otras palabras, contrariamente a nosotros no aprenden incrementalmente ni pueden relacionar lo nuevamente aprendido con lo que ya sabíamos. Además nosotros no necesitamos ver millones de ejemplos de gatos y perros para distinguirlos, con unos pocos basta. A medio plazo es posible conseguir desarrollar sistemas inteligentes más generalistas, es decir, no limitados como ahora a resolver una única tarea, sino capaces de ejecutar con excelencia varias a la vez y muy posiblemente serán sistemas que combinarán componentes de aprendizaje basado en el análisis de datos con componentes de razonamiento basados en conocimientos representados mediante lenguajes de representación basados en la lógica matemática y sus extensiones.

¿Cuál es pues el motivo por el que muchos creen que la IA está a punto de igualar a la inteligencia humana y, a partir de esta falsa premisa, hacen predicciones acerca de una posible singularidad tecnológica? En mi opinión, el excesivo antropocentrismo es el principal motivo de que la sociedad tenga una percepción errónea del estado real de la inteligencia artificial. Cuando nos informan de logros espectaculares de una IA específica en una competencia muy compleja aunque sea muy concreta tendemos a generalizar y atribuimos a la IA la capacidad de hacer prácticamente cualquier cosa que hacemos los seres humanos e incluso de hacerlo mucho mejor. En otras palabras, creemos que la IA prácticamente no tiene límites cuando de hecho es extremadamente limitada y, lo que es muy importante, no tiene casi nada que ver con la inteligencia humana, en realidad lo que tienen los actuales sistemas de IA no es inteligencia si no “competencia sin comprensión” en el sentido que apunta Daniel Dennett en su último libro “From bacteria to Bach and back”. Es decir, sistemas que pueden llegar a ser muy competentes llevando a cabo tareas específicas como discriminar una serie de elementos en una imagen pero sin comprender absolutamente nada acerca de la naturaleza de tales elementos ni de las propiedades y relaciones entre ellos debido a la ausencia de sentido común. Por ejemplo, pueden identificar una persona frente a una pared pero no saben que las personas no pueden atravesar paredes ni que las personas no pueden estar en dos lugares al mismo tiempo.

Es cierto que la inteligencia humana es el referente principal de cara a alcanzar el objetivo último de la IA, es decir la IA general comparable o incluso superior a la inteligencia humana, pero en mi opinión por muy sofisticada que llegue a ser la IA siempre será distinta de la humana ya que el desarrollo mental que requiere toda inteligencia compleja depende de las interacciones con el entorno y estas interacciones dependen a su vez del cuerpo, en particular del sistema perceptivo y del sistema motor. Ello, junto el hecho de que las máquinas muy probablemente no seguirán procesos de socialización y culturización, incide todavía más en el hecho de que, por muy sofisticadas que lleguen a ser, serán inteligencias distintas a las nuestras. El hecho de ser inteligencias ajenas a la humana y por lo tanto ajenas a los valores y necesidades humanas nos debería hacer reflexionar sobre posibles limitaciones éticas al desarrollo de la Inteligencia Artificial. En particular opino que ninguna máquina debería nunca tomar decisiones de forma completamente autónoma o dar consejos que requieran, entre otras cosas, de

la sabiduría, producto de experiencias humanas, así como de tener en cuenta valores humanos. Es decir que el peligro de la IA no es la singularidad tecnológica debida a la existencia de unas futuras hipotéticas súper-inteligencias artificiales, los verdaderos problemas ya están aquí y tienen que ver con la privacidad (vigilancia y control masivo de la ciudadanía, como por ejemplo el score social que se está implantando en China), la autonomía de los sistemas (por ejemplo las armas autónomas), la excesiva confianza acerca de sus capacidades (que es la razón por la que muchos hablan de substitución de personas por máquinas en prácticamente cualquier puesto de trabajo), el sesgo de los algoritmos de aprendizaje (con las consiguientes decisiones erróneas que ello supone) y la imposibilidad para rendir cuentas justificando sus decisiones en un lenguaje comprensible para las personas.

Consideremos para empezar el problema de la privacidad: Actualmente los algoritmos en los que se basan los motores de búsqueda en Internet, los sistemas de recomendación y los asistentes personales de nuestros teléfonos móviles, conocen bastante bien lo que hacemos, nuestras preferencias y nuestros gustos e incluso pueden llegar a inferir lo que pensamos y como nos sentimos. El acceso a cantidades masivas de información, que voluntariamente generamos, es fundamental para que esto sea posible ya que mediante el análisis de estos datos provenientes de fuentes diversas es posible encontrar relaciones y patrones que serían imposibles de detectar sin las técnicas de IA. Todo ello resulta en una pérdida alarmante de privacidad. Para tratar de evitarlo deberíamos tener derecho a poseer una copia de todos los datos personales que generamos, controlar su uso y decidir a quién permitimos el acceso a los mismos y bajo qué condiciones en lugar de que estén en manos de grandes corporaciones.

Sigamos con el problema de la autonomía: La IA está basada en programación compleja, y por lo tanto necesariamente cometerá errores. Pero incluso suponiendo que fuera posible desarrollar software completamente fiable, hay dilemas éticos que los desarrolladores de software deberán tener en cuenta a la hora de diseñar sistemas autónomos. Por ejemplo, un vehículo autónomo podría decidir atropellar a un peatón para evitar una colisión que podría causar daños a sus ocupantes. Otro ejemplo claro son las armas autónomas. Los tres principios básicos que rigen los conflictos armados: discriminación (la necesidad de discernir entre combatientes y civiles o entre un combatiente rindiéndose y uno en disposición de atacar), proporcionalidad (hasta qué punto son aceptables los daños colaterales) y precaución (minimización del número de víctimas) son extraordinariamente difíciles evaluar, y por consiguiente casi imposible de cumplir por los sistemas de IA que controlan las armas autónomas. Pero incluso en el caso en que a muy largo plazo las máquinas tuvieran estas capacidades, en mi opinión sería indigno delegar en una máquina la decisión de matar. En general, cuanto más autonomía demos a los sistemas de IA más responsabilidad deberíamos exigir a los diseñadores y programadores de dichos sistemas de tal forma que cumplan principios legales y éticos. Es decir, el verdadero problema no es el monstruo de Frankenstein sino Frankenstein.

En cuanto al impacto en el trabajo, a pesar de ser cierto que el impacto en el empleo será importante, posiblemente no lo sea tanto como algunos

prevén. Sin duda, el innegable entusiasmo actual por la IA nos puede hacer creer que la inteligencia humana es sustituible, y ello ha llevado a algunas organizaciones a despedir empleados y reemplazarlos por sistemas de IA. Esto es un error grave ya que de hecho todos los sistemas de IA dependen críticamente de la inteligencia humana. Los sistemas basados en el conocimiento se basan en el conocimiento y la comprensión de la experiencia humana, y los sistemas de IA que se basan en datos dependen críticamente de datos acerca de la conducta humana. De aquí se desprende que hay que continuar enseñando, desarrollando y ejerciendo la capacidad humana. Por otro lado, en la gran mayoría de casos, la capacidad humana todavía supera con creces la inteligencia artificial, especialmente cuando el sistema de IA debe enfrentarse a situaciones que no han aparecido en los conjuntos de datos con los que se han entrenado los sistemas de IA. Por otra parte, a menudo, muchas aplicaciones se benefician de la sinergia entre el ser humano y la inteligencia artificial, es decir que la unión persona-máquina está produciendo resultados superiores a cualquiera de los dos por separado ya que por muchos datos que pueda analizar una máquina siempre será necesario el juicio humano, uno de los motivos es que las máquinas no pueden distinguir entre correlación y causalidad. Este fenómeno se está dando en ámbitos como el diagnóstico médico y la toma de decisiones en general incluidas las decisiones empresariales. Los trabajadores necesitarán recibir una formación continuada que les permita adaptarse a nuevas formas de trabajo que requerirán más creatividad, colaboración entre ellos y con las máquinas, e iniciativa para puestos de trabajo cambiantes organizados por tareas concretas con una elevada movilidad tanto geográfica como funcional. Las empresas por su parte deberán invertir mucho más en IA y en particular en la formación continua de sus empleados incluyendo los ejecutivos. Todavía pocas empresas han incorporado la IA en su cadena de valor, uno de los principales factores limitativos es la escasez de personas con una formación adecuada en IA.

Otro problema son los sesgos de los algoritmos. Ningún sistema de IA tiene intencionalidad pero las decisiones que toman están basadas en datos de entrenamiento que a menudo están sesgados con lo que las decisiones que toman están también sesgadas. La IA no solamente reproduce los sesgos humanos sino que los amplifica. Por ejemplo, un sistema de pre-selección de candidatos a un puesto de trabajo de nivel directivo fue entrenado con datos históricos que reflejaban estadísticamente que los ejecutivos más exitosos eran hombres blancos por lo que el sistema discriminaba a candidatos mujeres y afroamericanos, el problema es que el algoritmo no tenía en cuenta la minoritaria presencia de mujeres y personas no blancas en los datos de entrenamiento del algoritmo. Otro mal ejemplo es un sistema de análisis de imágenes que después de ser entrenado con miles de imágenes cotidianas asoció las imágenes de mujeres con imágenes de cocinas pero imágenes de hombres con actividades deportivas. Otro ejemplo si cabe más preocupante, es el software COMPASS usado por jueces en EEUU para evaluar la probabilidad de reincidencia que otorgaba una probabilidad doblemente superior de reincidir a ciudadanos afroamericanos que a ciudadanos blancos. Este sesgo fue detectado y denunciado y tuvo que ser corregido pero ya había influido en muchas decisiones de jueces antes de su corrección. Es necesario establecer metodologías de verificación y validación adecuadas de los algoritmos de IA con el fin de que sean utilizadas por autoridades certificadoras, algo así como

los procesos de certificación de la seguridad de los alimentos que consumimos o los medicamentos que tomamos.

El último problema que quería mencionar es la rendición de cuentas de los algoritmos. Cuando un sistema de IA toma decisiones, las personas afectadas por estas decisiones tienen que poder recibir una explicación de por qué se toma la decisión en un lenguaje comprensible y tienen que ser capaces de cuestionar la decisión con argumentos razonados. Esto es especialmente importante en campos como, por ejemplo, decisiones sobre préstamos, decisiones legales (por ejemplo, en la concesión de una libertad condicional), seguros, impuestos, etc. Muchos sistemas de inteligencia artificial, en particular los que se basan en enfoques a partir de datos, actualmente no pueden proporcionar este tipo de explicación. Sus decisiones se derivan de un amplio conjunto de parámetros obtenidos estadísticamente. Estamos en los inicios de las investigaciones sobre técnicas para comprender el funcionamiento de estos sistemas, y es probable que, de nuevo, sea necesaria una combinación de IA basada en el conocimiento y de IA basada en los datos. La rendición de cuentas es, claramente, una condición previa a cualquier despliegue racional de aplicaciones de la IA.

En definitiva, la IA y sus algoritmos no son neutrales, sino el reflejo de las intenciones y los sesgos del equipo de programadores y entidades implicados en su implementación, con el añadido de que no solo los reflejan sino que los amplifican. Todos estos problemas relacionados con el impacto de la IA, hacen que muchos expertos señalemos la necesidad de regular su desarrollo.

Pero, además de regular, es imprescindible educar a los ciudadanos sobre los beneficios y riesgos de las tecnologías inteligentes (que no son los que vemos en las películas de ciencia ficción), dotando (a los ciudadanos) de las competencias necesarias para controlarla en lugar de ser controlados por ella. Necesitamos futuros ciudadanos mucho más informados, con más capacidad para evaluar los riesgos tecnológicos, con mucho más sentido crítico y capaces de hacer valer sus derechos. Este proceso de formación debe empezar en las escuelas y tener continuación en la universidad. En particular es necesario que los estudiantes de ciencia e ingeniería, reciban una formación ética que les permita comprender mejor las implicaciones sociales de las tecnologías que van a desarrollar. Solo si invertimos en educación conseguiremos una sociedad que pueda aprovechar las ventajas de las tecnologías inteligentes minimizando sus riesgos y de esta forma la Inteligencia Artificial servirá para dar un gran paso en el progreso de la humanidad.

Referencias:

[1] A. Ribes, J. Cerquides, Y. Demiris, R. Lopez de Mantaras; Active Learning of Object and Body Models with Time Constraints on a Humanoid Robot. **IEEE Transactions on Cognitive and Developmental Systems** 8(1): 26-41, 2016

Ramon López de Mántaras Badia

Doctor en Informática.

Profesor de Investigación del CSIC y Director del Instituto de Investigación en Inteligencia Artificial.

Capítulo 17

Encontrar vida en otros mundos: un reto futuro para uno de los sueños más antiguos de la humanidad.

Nuestra relación con los cuerpos celestes.

La humanidad, desde sus orígenes, ha sentido una gran fascinación por el cielo y esas luces brillantes que aparecían suspendidas en él y que parecían regir nuestros destinos. El Sol y la lluvia hacían prosperar las cosechas, pero las tormentas y las sequías también podían destruirlas. El Sol nos iluminaba de día y la Luna de noche, aunque a veces sucedieran fenómenos extraños y, repentinamente, el planeta se quedara sumido en una gran oscuridad. Nuestro ritmo cotidiano parecía marcado por los astros, así que no es extraño que pensáramos que estos gobernaban nuestras vidas y los identificáramos con dioses, a los que había que rendir tributo para que fueran benévolos. En el antiguo Egipto se adoraba a Ra, en Grecia a Helios y en Persia a Mitra, nombres todos ellos que designaban al mismo dios: el dios del Sol. No debería extrañarnos, por tanto, que en muchas civilizaciones antiguas se consultaran las estrellas para predecir el futuro. Aun hoy, mucha gente sigue creyendo que el horóscopo puede ayudarles a saber qué les va a deparar el día a día y cómo pueden actuar mejor para resolver los retos de la vida.

A pesar de esta mezcla de temor y veneración por los cuerpos celestes, la humanidad, durante gran parte de su historia, mantuvo muy arraigada la creencia de que la Tierra era el centro del Universo, alrededor del cual giraban el resto de planetas del Sistema Solar. Al mismo tiempo, la mayoría de las religiones y mitologías apoyaban la idea de que el hombre era el rey de la creación, al que estaba supeditada el resto de la vida terrestre. En ese contexto, la Tierra se nos mostraba como un planeta especial que, muy probablemente, era el único habitado. En el hipotético caso de que existiera vida en otros mundos, esta sería obra del mismo Creador que había originado la vida terrestre, así que no debería ser muy diferente de ella. En ese escenario el hombre seguía siendo la especie destinada a reinar sobre todas las demás que pudieran poblar el Universo.

Esta visión cambió radicalmente cuando, en el siglo XVI, Nicolás Copérnico formuló la teoría heliocéntrica, demostrada unas décadas más tarde por Galileo Galilei, la cual afirmaba que el centro del Universo conocido no era la Tierra, sino el Sol. A partir de ese momento, tímidamente primero y más abiertamente después, comenzó a especularse con la posibilidad de que nuestro planeta no fuera tan especial como nos creíamos y, por tanto, no sería inconcebible que pudiera haber otros mundos similares, también poblados de vida.

El punto de inflexión respecto a la consideración del lugar ocupado por el hombre en el conjunto de la vida terrestre llegó a mediados del siglo XIX, cuando Charles Darwin publicó su libro *El origen de las especies*, en el que exponía sus teorías sobre la evolución y la diversificación de la vida. Años más tarde, el descubrimiento del ADN como molécula de la herencia y la constatación de las similitudes que a nivel molecular comparten los seres vivos condujeron a la conclusión de que toda la vida de la Tierra procede de un ancestro común. Han sido los procesos de mutación y selección natural, actuando de forma concertada a lo largo de miles de millones de años, los que han dado lugar a toda la diversidad de la vida a partir de ese remoto antepasado. Esta visión despoja al ser humano del papel que creía tener como rey de la creación. En lugar de eso, parece que somos una especie más, producto del azar y de las características ambientales por las que ha pasado nuestro planeta a lo largo de su historia.

Cuando quedó claro que la Tierra era un planeta más entre otros muchos, y el hombre una especie más entre todas las que existen, desaparecieron muchos de los prejuicios mentales que nos impedían aceptar que pudiera haber vida en otros planetas y que, en caso de haberla, pudiera ser diferente de la vida que conocemos. Poco a poco, el desarrollo tecnológico y los avances científicos nos han permitido explorar los mundos más cercanos al nuestro y vislumbrar lo insignificantes que somos en un Universo que parece estar plagado de planetas. Actualmente estamos en condiciones de llegar a regiones muy alejadas de nuestro Sistema Solar y de poder investigar la presencia de vida en ellas. Pero antes de hablar de todo esto, merece la pena detenernos un poco para destacar tres grandes hitos científicos que han aumentado considerablemente nuestras expectativas de encontrar vida fuera de la Tierra.

La vida puede ser un fenómeno común en el Universo.

El primer gran hito que cambió nuestra concepción de lo que era la vida y como podría ser de frecuente en el Universo tuvo lugar en 1953, fecha en la que el químico Stanley Miller realizó su famoso experimento citado en multitud de libros de texto. Este consistió en mezclar en una cámara los gases que se creía que componían la atmósfera de la Tierra primitiva (metano, amoníaco, vapor de agua e hidrógeno) y aplicar una descarga eléctrica que proporcionara la energía necesaria para que dichos gases reaccionaran. Como resultado, se sintetizaron diversos aminoácidos, algunos azúcares y varias de las bases nitrogenadas que forman parte de los ácidos nucleicos, demostrándose así que muchos de los componentes esenciales de la materia viva podían ser sintetizados en condiciones abióticas. El experimento de Miller marcó el inicio de un nuevo campo de investigación, la química prebiótica, enfocado a investigar la síntesis de compuestos orgánicos en las condiciones físicas y con los ingredientes químicos que pudieron estar presentes en la Tierra primitiva. Actualmente ha sido posible sintetizar en el laboratorio la mayoría de los componentes básicos de la vida. Aunque no podemos afirmar que las reacciones químicas seguidas para la síntesis de estos compuestos sean las que precedieron al origen de la vida en la Tierra, los hallazgos tienen una

importancia crucial, ya que demuestran que, si condiciones similares pudieran darse en otros planetas, ¿por qué no podría surgir la vida en ellos?

El segundo gran hito que queremos reseñar es el reconocimiento de que la vida es robusta y puede soportar condiciones físico-químicas que antes eran consideradas incompatibles con su existencia. El descubrimiento de los microorganismos extremófilos que, no solo soportan condiciones extremas, sino que incluso las requieren, ha cambiado enormemente nuestra idea sobre cómo deben de ser los lugares en los que puede prosperar la vida. Actualmente conocemos microorganismos termófilos (capaces de vivir a temperaturas mayores de 100°C), acidófilos (que necesitan pHs muy ácidos o muy básicos), halófilos (que viven en presencia de concentraciones muy

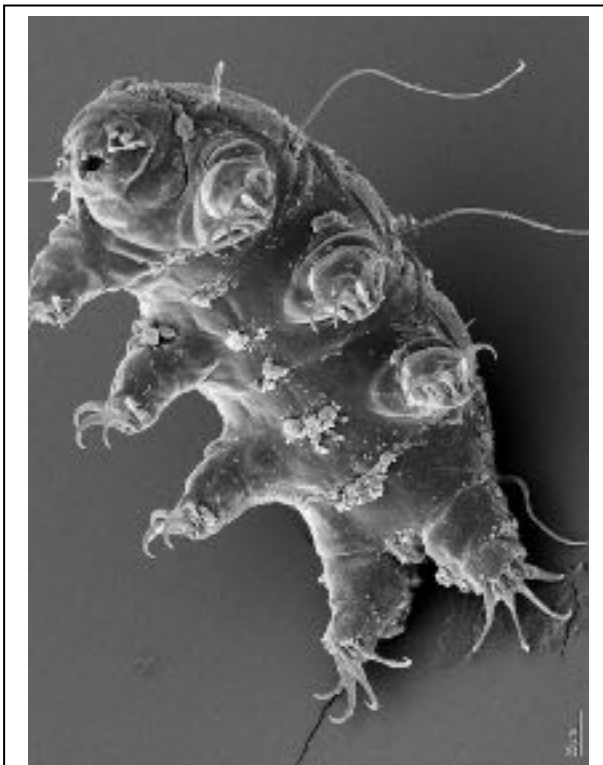


Fig.1. Un tardígrado visto al microscopio electrónico. Créditos: ESA/Dr. Ralph O. Schill.

elevadas de sal), barófilos (que soportan presiones mucho más elevadas que las que existen en la superficie terrestre)... incluso existen unos pequeños animales invertebrados, los tardígrados, los cuales, ante circunstancias adversas, son capaces de deshidratarse casi completamente, pudiendo permanecer varios años en ese estado. Los tardígrados, además, pueden sobrevivir a temperaturas que oscilan entre los -272°C y los 150°C, resistir presiones seis veces superiores a las del más profundo de los océanos, radiaciones ionizantes, la inmersión en alcohol puro o éter y la exposición a condiciones de vacío. Todas estas formas de vida “extrema”, que podemos encontrar sin salir de nuestro planeta, quizás sean solo algunos ejemplos del tipo de vida que podemos esperar encontrar fuera de él.

Por último, el tercer gran hito que mencionaremos es el descubrimiento de la gran cantidad de planetas extrasolares (que orbitan alrededor de una estrella que no es el Sol) que parece existir en el Universo. En diciembre de 2018, la NASA había confirmado la existencia de casi 4.000 de esos planetas, muchos de los cuales son planetas rocosos, como la Tierra, y están situados en la zona de habitabilidad de su estrella. Actualmente no podemos saber cómo es la atmósfera de esos planetas y si podrían albergar vida o no, algo que es esperable que las nuevas generaciones de telescopios nos ayuden a determinar. Sin embargo, lo más sorprendente es que los planetas extrasolares identificados hasta ahora son solo una mínima parte de los cientos de trillones que se cree que puede haber en el Universo. ¡Con estos números es muy difícil pensar que en alguno de ellos no pueda existir la vida!



Fig. 2. Recreación artística de la superficie del exoplaneta TRAPPIST-1f, localizado en el sistema TRAPPIST-1 en la constelación de Acuario. Créditos: NASA/JPL-Caltech.

Nuestra guía en la búsqueda de vida extraterrestre.

Uno de los grandes problemas que tenemos a la hora de buscar vida fuera de la Tierra es lo difícil que es encontrar una aproximación que no esté sesgada por las propiedades de la vida que conocemos, un hecho que podría hacernos descartar las formas que fueran muy diferentes a la vida terrestre. Es necesario, por tanto, definir qué características debería tener una estructura que encontráramos en otro planeta para poder afirmar con certeza que está viva. Esta cuestión nos lleva a examinar en detalle la vida terrestre, para intentar así discernir cuáles de sus propiedades son esenciales y cuáles soluciones particulares adoptadas en el ambiente químico-físico particular de nuestro planeta.

Una primera aproximación para determinar qué es lo esencial de la vida puede consistir en tratar de identificar qué es lo que tienen en común las distintas formas en que esta se manifiesta. Si iniciamos este ejercicio, muy pronto nos daremos cuenta de que, a pesar de las grandes diferencias que presentan las especies terrestres, a nivel molecular todas ellas son muy parecidas. Son tan parecidas que ahora sabemos que tan alto nivel de coincidencia solo es posible si tienen el mismo origen, lo que complica enormemente saber cuáles de las características de la vida terrestre son esenciales y cuáles, simplemente, han sido heredadas de ese progenitor común.

- Comencemos por el **metabolismo**: la transformación de la materia y la energía externas en materia y energía utilizables por la vida. Todos los seres vivos, sin excepción, realizan este proceso, el cual implica la realización de un gran número de reacciones químicas de forma coordinada y precisa que son catalizadas por las proteínas con actividad enzimática. Las proteínas son largas

cadena formada por la unión de unos monómeros denominados aminoácidos que, cuando están en un solvente acuoso, se pliegan formando estructuras tridimensionales que son las que confieren mayor estabilidad a la molécula y determinan cuál es su función específica. Los aminoácidos pueden presentarse en dos formas que son imágenes especulares la una de la otra. Estas dos formas, que se denominan L y D se forman con igual facilidad y pudieron ser igualmente abundantes en la Tierra primitiva. Sin embargo, las proteínas biológicas solo contienen aminoácidos L. Intrigante, ¿verdad?

- Otra propiedad de la vida es la necesidad de contener una **información** que permita saber el orden en que se tienen que disponer los aminoácidos para que se formen las proteínas correctas en el momento adecuado. Esta información es la información genética, que se almacena en una molécula, el ácido desoxirribonucleico o ADN, utilizando un código que es universal. El código genético está basado en el orden en que se disponen las unidades básicas – los nucleótidos- que forman la molécula de ADN. Cada secuencia de tres nucleótidos en el ADN (lo que se denomina codones o tripletes) especifica la incorporación de un aminoácido concreto en las proteínas. Y esto es así en todos los seres vivos, sin que todavía se haya encontrado una razón válida para esa coincidencia.



Fig.3. Recreación artística de la molécula de ADN. Fuente: Pixabay.

- La última propiedad de la vida que citaremos es el **aislamiento del entorno**. Es fácil imaginar que los componentes de la vida tienen que estar reunidos en un compartimento, de modo que se evite su dilución en el medio y se facilite su funcionamiento integrado. En todas las formas de la vida terrestre esos compartimentos son las células, separadas del medio externo por las membranas plasmáticas, unas complejas estructuras capaces de cumplir con la función de aislamiento sin interrumpir la comunicación con el exterior.

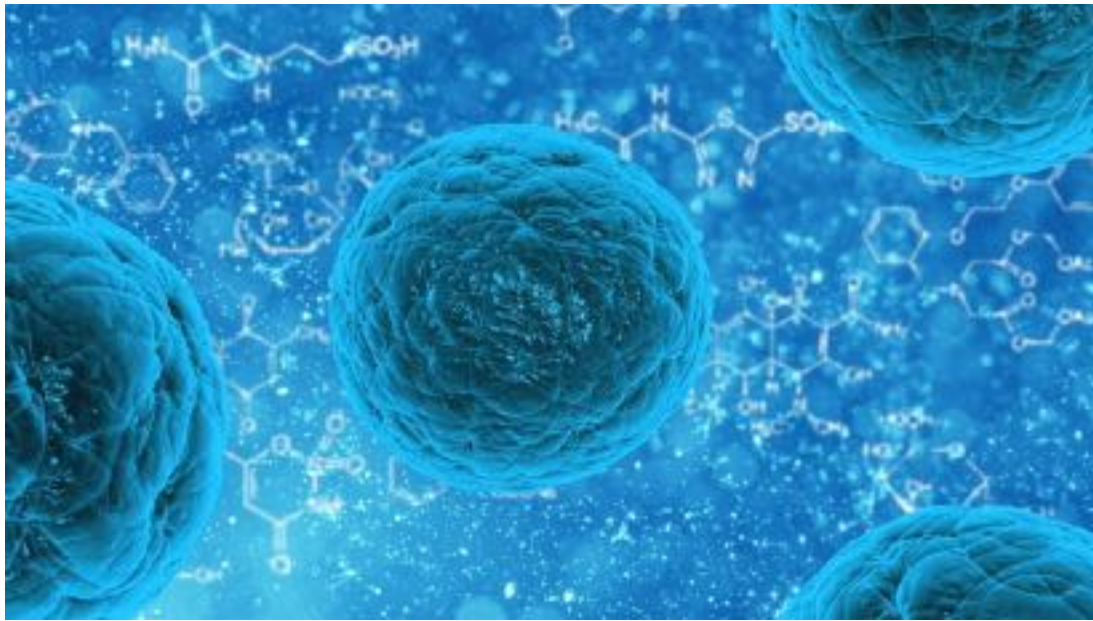


Fig.4. Células. Fuente: Public Domain Pictures.

Se puede pensar que las tres grandes propiedades que hemos mencionado: metabolismo, información y compartimento son esenciales para la vida y, por tanto, nuestra búsqueda de vida en otros planetas debería estar orientada a la búsqueda de sistemas que las posean. Sin embargo, ¿deberíamos restringirnos a la búsqueda de las moléculas y estructuras que conforman la vida terrestre? No es descabellado pensar que alguna forma de vida extraterrestre pudiera utilizar para su metabolismo fuentes de energía que no utiliza la vida de la Tierra, como la energía de la gravedad, la del viento, la de las mareas o la calorífica. También sería posible imaginar una catálisis más básica, que no implicara moléculas tan complejas como las proteínas, o que estas estuvieran formadas por la forma D de los aminoácidos en lugar de por la L. La información podría almacenarse en moléculas diferentes de los ácidos nucleicos o emplear un código distinto. Incluso, la vida podría estar basada en un elemento químico diferente del carbono. No debemos olvidar que el ambiente físico-químico imperante en cada planeta es el que va a condicionar las reacciones que en él ocurren y, por tanto, el tipo de vida que podría surgir.

La definición de vida más aceptada actualmente es la propuesta por la NASA, la cual propone que la vida es un sistema químico automantenido capaz de evolución darwiniana. Esta concisa definición desliga la vida de un ambiente particular, incluye la necesidad de un metabolismo y, también, de una información para que ese metabolismo pueda funcionar. Y algo más: considera la capacidad de evolución darwiniana, algo que surge como consecuencia de que la información debe transmitirse a las generaciones siguientes, para lo cual necesita ser copiada. Durante la copia inevitablemente se producen errores, que darán lugar a diferencias en las características de los organismos, permitiendo así la acción de la selección natural y, con ello, la evolución de la vida.

Una definición como la que acabamos de describir amplía enormemente el tipo de vida que podemos esperar encontrar fuera de la Tierra, pero al mismo tiempo dificulta enormemente su identificación, al no estar restringida a un tipo de moléculas o estructuras concretas. ¡Necesitamos tener la mente muy abierta para aceptar que la vida quizás pueda manifestarse en formas que no alcanzamos siquiera a imaginar!

¿Qué se necesita para que pueda surgir la vida?

Hasta hace no mucho tiempo se consideraba que para que pudiera surgir la vida en un planeta, este debería presentar unos valores para sus variables físico-químicas similares a los que parecen más favorables para la vida en la Tierra. Sin embargo, como hemos visto, esta idea ha cambiado radicalmente con el descubrimiento de los microorganismos extremófilos y actualmente se acepta que las condiciones que requiere la vida son únicamente la presencia de agua líquida, la existencia de elementos químicos con los que poder construir las moléculas biológicas y la accesibilidad a alguna fuente de energía que pueda ser empleada en el metabolismo.

La presencia de agua líquida parece una condición casi ineludible, dadas sus excelentes propiedades como solvente que la convierten en uno de los mejores medios para que las reacciones químicas del metabolismo puedan tener lugar. Los elementos químicos fundamentales que componen la vida de la Tierra son carbono, hidrógeno, oxígeno, nitrógeno y azufre. El carbono forma los esqueletos básicos de las moléculas biológicas, y el resto de elementos pueden formar grupos funcionales que facilitan la realización de reacciones químicas y la construcción de moléculas complejas a partir de unidades más simples. Por último, la existencia de fuentes de energía es algo esencial para el metabolismo, pero, como ya hemos mencionado, no deberíamos restringirnos a las que emplea la vida terrestre. Hasta hace no mucho, se pensaba que toda la energía utilizada por la vida terrestre procedía en último término de la radiación solar, pero se ha visto que puede haber ecosistemas completos que funcionan en completa oscuridad, como los que se han encontrado en las cercanías de las chimeneas volcánicas submarinas y en varios lugares del subsuelo, un ambiente que solo recientemente hemos comenzado a explorar. La fuente de energía primaria en estos ecosistemas es la desprendida en algunas reacciones químicas que se producen entre ciertos minerales de las rocas, algo que, como veremos, puede no ser raro en otros lugares del Universo.

La importancia del agua para la vida es tal que ha llevado a definir la zona de habitabilidad de una estrella como la región alrededor de la misma en la que sus planetas asociados podrían mantener agua líquida de forma estable sobre su superficie. Sin embargo, los hallazgos sobre la existencia de vida en el subsuelo terrestre deberían ampliar esta definición, para incluir la posibilidad de que el agua líquida pueda estar también bajo la superficie.

¿Qué posibilidades hay de encontrar vida en nuestro Sistema Solar?

Los lugares del Sistema Solar que ofrecen más expectativas de poder albergar vida son: Marte, Europa (la luna de Júpiter) y dos de los satélites de Saturno (Encélado y Titán).

Marte es un planeta rocoso que se formó al mismo tiempo que la Tierra, hace aproximadamente 4.600 millones de años. Inicialmente ambos planetas debieron compartir características ambientales similares, como la presencia en su atmósfera de gran cantidad de gases de origen volcánico que generarían un efecto invernadero suficiente para que el clima fuera templado, favoreciéndose así la existencia de agua líquida en la superficie. Las imágenes obtenidas por los orbitadores que sobrevuelan la superficie de Marte muestran en la zona sur del planeta lo que parecen las huellas de gigantescas redes fluviales, así como los valles excavados por ellas. La presencia de algunos minerales como arcillas, hematita y jarosita, que normalmente se forman en condiciones de humedad, no hace sino añadir pruebas a la existencia de un pasado en Marte con abundante agua líquida fluyendo por su superficie. En esta época también parece que Marte poseía un campo magnético que le protegía del viento solar y de la radiación cósmica. La pregunta que surge entonces es: ¿qué pasó para que Marte evolucionara de forma tan distinta a como lo ha hecho la Tierra?



Fig. 5. Evidencias geológicas del flujo de agua líquida en la superficie de Marte.

Créditos: NASA/JPL-Caltech/University of Arizona.

Hace aproximadamente 4.000 millones de años, conforme el interior de Marte se iba enfriando, las emisiones de los volcanes comenzaron a decrecer y, por alguna razón que desconocemos, Marte perdió su campo magnético. Sin esa protección, la atmósfera casi desapareció, barrida por el viento solar. Y, con la desaparición de la atmósfera, el planeta se enfrió considerablemente, hasta convertirse en el árido lugar que es actualmente. La temperatura media de Marte es de -55°C . En verano, en la zona del Ecuador, las temperaturas máximas durante el día pueden llegar a los 20°C , pero no es raro que las mínimas nocturnas desciendan hasta los -80°C . La temperatura en invierno en los polos es todavía menos acogedora, ya que puede llegar a alcanzar los -143°C .

El bajo espesor de la atmósfera de Marte, unido a sus bajas temperaturas, provoca que el agua no se pueda mantener en forma líquida en su superficie. En su lugar, se encuentra en forma de hielo y, cuando las temperaturas se elevan en verano, pasa directamente al estado de gas. En estas condiciones no solo se congela el agua, sino también el dióxido de carbono atmosférico, formando hielo que recubre la mayor parte del planeta y

que forma grandes masas en los polos. Otros problemas para la vida derivados del hecho de tener una atmósfera tan tenue es que apenas ofrece protección frente a los rayos ultravioleta y otras radiaciones peligrosas procedentes del Sol. Peor aún que la radiación ultravioleta es la radiación cósmica procedente del espacio, frente a la cual Marte tampoco tiene protección debido a la ausencia de campo magnético. Por si todo lo anterior no bastara, Marte sufre tormentas de polvo a escala planetaria, con vientos que pueden llegar a alcanzar los cien kilómetros por hora que remueven y llevan de un lugar a otro la capa de polvo rojo que recubre todo el planeta.

Según los datos enviados por la sonda Mars Odissey, parte del agua que había en el Marte primitivo permanece congelada en los casquetes polares, sobre todo en el del Polo Norte. Respecto al resto del agua, hay varias posibilidades no excluyentes: que escapara al espacio, que migrara al interior del planeta o que se encuentre congelada entre los granos de las rocas, formando un tipo de suelos similar al permafrost terrestre. Recientemente, en julio de 2018, un equipo de científicos italianos ha detectado lo que parece ser un gran lago de agua líquida salada bajo el hielo del Polo Sur de Marte. Actualmente no podemos saber si se trata de agua líquida pura o de rocas porosas impregnadas de agua. Su temperatura parece ser muy inferior a los cero grados, pero probablemente contiene una elevada concentración de sales que podrían actuar como anticongelante. Tampoco es descartable la existencia de alguna fuente de calor interno que, ayudada por la mayor presión que existe en el subsuelo, contribuya a mantener el agua en estado líquido. La recién aterrizada misión InSight, destinada a explorar el interior de Marte, quizás pueda decirnos como de probable es que existan esas fuentes de calor.

Todo lo que hemos descrito indica que, aunque las condiciones actuales de la superficie marciana no parecen ser muy aptas para la vida, sí podrían serlo en el subsuelo. Una capa de roca o polvo de unos milímetros de espesor proporciona suficiente protección frente a la radiación ultravioleta, pero se requiere al menos un metro de profundidad para bloquear los rayos cósmicos. Si en el subsuelo pudieran tener lugar reacciones químicas en las que se desprendiera energía, no sería descabellado pensar que esa energía pudiera acoplarse a la realización de un metabolismo por algún ser vivo similar a algunas bacterias terrestres. La gran masa de agua subterránea de la que acabamos de hablar podría ser uno de esos ambientes favorables para la vida.

Otras observaciones, como la presencia de metano en la atmósfera marciana, también han contribuido a alimentar las expectativas de que Marte pueda contener vida. El metano es un gas muy poco estable que se disocia muy fácilmente por acción de la luz solar. Por tanto, su permanencia en la atmósfera de un planeta indica que debe haber una fuente continua de emisión. En la Tierra existen un tipo de bacterias, las bacterias metanógenas, que liberan metano como un subproducto de su metabolismo, lo que ha hecho que se especule con la existencia en Marte de bacterias similares. Sin embargo, existen alternativas geoquímicas que pueden explicar el origen del metano marciano sin implicar la existencia de vida, lo que hace que la presencia de este gas no pueda ser considerada una evidencia de vida en Marte.

La misión ExoMars 2020 de la ESA llegará a Marte muy pronto, en el año 2020. Su principal objetivo será buscar signos de vida, actual o pasada. Será la primera misión en extraer y analizar muestras del subsuelo de Marte, donde ya hemos visto que las condiciones para que pueda haber vida son más favorables que en la superficie. ¡Quién sabe qué sorpresas nos encontraremos!

Europa, el satélite de Júpiter, es un mundo helado, algo menor que nuestra Luna. Está compuesto principalmente por silicatos y cubierto por una capa de hielo de aproximadamente diez kilómetros de espesor que aparece surcada por numerosas grietas y rayas. Europa posee un campo magnético, que probablemente está inducido por el campo magnético de Júpiter, lo que implica la existencia de una capa conductora a no más de doscientos kilómetros de profundidad. Actualmente sabemos que esa capa es un gigantesco océano subterráneo de agua salada de aproximadamente cien kilómetros de espesor que se mantendría líquido gracias a la energía de las mareas causadas por la atracción de Júpiter. El hecho de que Europa posea agua líquida bajo su superficie ha alimentado las expectativas de que ese océano pudiera albergar vida, sobre todo si estuviera en contacto con el interior rocoso del planeta. Ese hecho facilitaría la disolución de minerales que podrían intervenir en reacciones químicas en las que se desprendiera energía que podría ser utilizada por la vida.

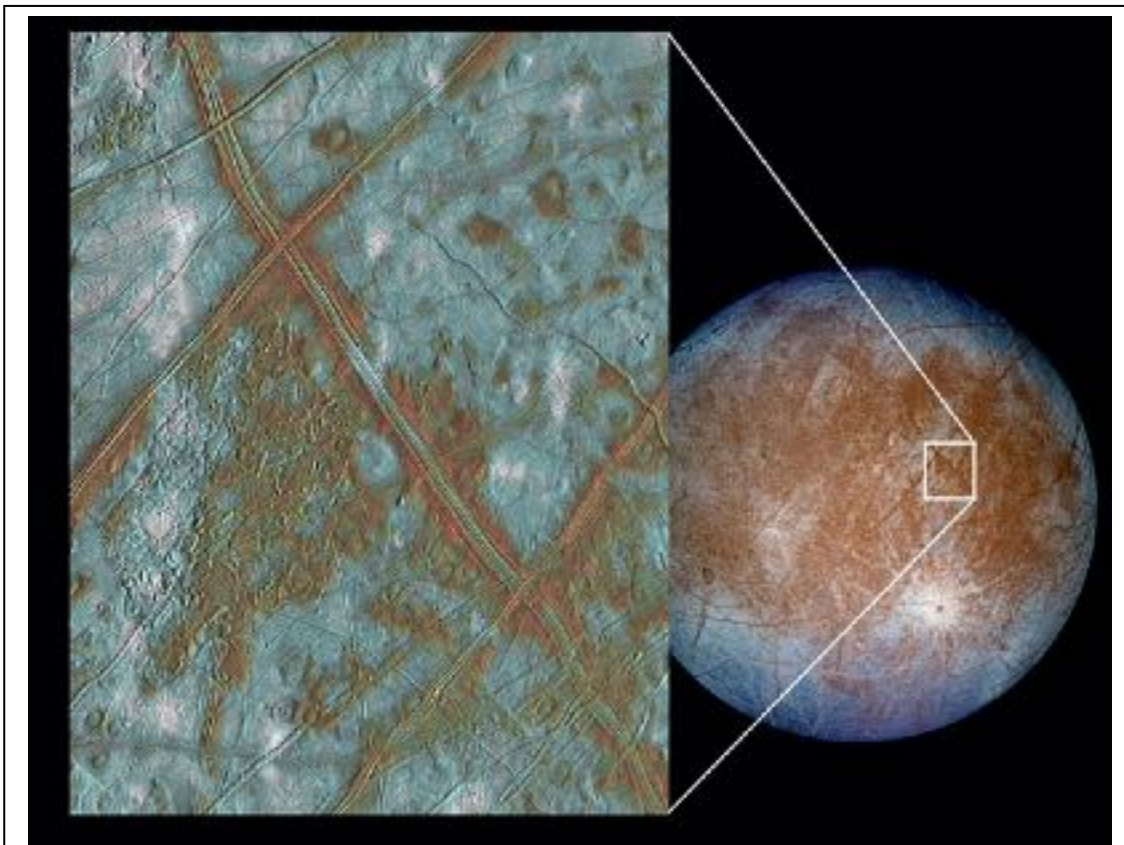


Fig.6. Imagen de Europa, el satélite de Júpiter, mostrando las fracturas en el hielo que lo surcan. Créditos: NASA/JPL/University of Arizona.

Las características **Encélado** son en gran medida similares a las de Europa, ya que este satélite de Saturno también posee un océano interno en el cual, además, parece existir importante actividad hidrotermal.

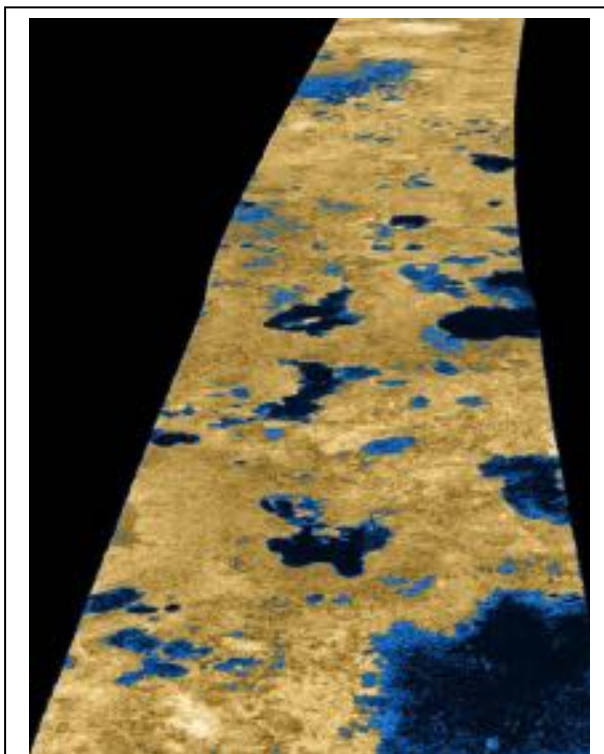


Fig.7. Lagos en la superficie de Titán. Créditos: NASA/JPL-Caltech/USGS.

Por último, hablaremos de **Titán**, el satélite más grande de Saturno, un mundo rico en moléculas orgánicas, y también considerado como potencialmente habitable. La atmósfera de Titán es muy densa. Está formada por un 98% de nitrógeno y un 2% de metano. La reacción de estos gases con la luz ultravioleta puede producir una amplia cantidad de moléculas orgánicas que posteriormente se depositarían en la superficie. Los modelos sobre la formación y la evolución de Titán predicen que en su interior hay una capa de hielo de agua, debajo de la cual hay un océano de agua líquida de varios cientos de kilómetros de espesor. Titán contiene lagos y mares compuestos de metano y etano en estado líquido, los cuáles podrían albergar vida basada en un

solvente distinto del agua. Sin embargo, las bajas temperaturas de Titán – alrededor de -180°C – representan un gran obstáculo para la realización de cualquier tipo de metabolismo que pudiera conducir a la generación de un elevado grado de complejidad biológica.

Aún no sabemos que podemos encontrar en estos satélites helados. En los próximos años está previsto que se lancen una serie de misiones para su exploración: Europa Clipper, Jupiter Icy Moons Explorer (JUICE), Dragonfly... ¡Confiamos que alguna de ellas nos depare el ansiado ejemplo de vida con un origen distinto de la terrestre!

¿Qué posibilidades hay de encontrar vida extraterrestre inteligente?

Todo lo que hemos expuesto hasta ahora nos indica que si encontramos vida en los planetas más cercanos a la Tierra, esta será probablemente microscópica. Eso puede resultar una decepción para las personas que lo que ansían es conocer si existen en el Cosmos otros seres que también estén dotados de inteligencia y con los que, quizás, podríamos comunicarnos. Ya las sondas Pioneer 10 y Pioneer 11, lanzadas en 1973, iban equipadas con una placa metálica que contenía un mensaje visual cuyo objetivo era informar de nuestra existencia a los seres de alguna posible civilización extraterrestre. Un

poco más adelante, en 1977, las sondas Voyager también contemplaron la posibilidad de comunicarse con otros habitantes del espacio y para eso incluyeron dos discos de cobre recubiertos de oro (uno en cada nave) que contenían grabada información sobre nuestra cultura. También en los años setenta, desde el observatorio de Arecibo en Puerto Rico se envió un mensaje de radio hacia el cúmulo globular M13, a 25.000 años luz de distancia de la Tierra. Las malas noticias son que, aunque obtuviéramos respuesta, esta tardaría más de 50.000 años en llegar.

El proyecto SETI (Search for Extraterrestrial Intelligence) es, desde la década de los 70, uno de los programas que más recursos ha invertido en la búsqueda de vida extraterrestre inteligente, o al menos con una inteligencia con capacidad para emitir señales electromagnéticas que puedan ser detectadas con un radiotelescopio. Hasta el momento no se ha encontrado ninguna señal que no pueda ser explicada por fenómenos astrofísicos naturales. Pero no por eso se ha dejado de rastrear el espacio a la espera de obtener un mensaje que nos haga vislumbrar la posibilidad de que existan otros seres que también estén buscando a sus compañeros en el Cosmos.

Hay científicos, como el recientemente fallecido Stephen Hawking, que opinan que debemos ser muy cautos con la información que enviamos al espacio sobre nosotros. En el interior de nuestra galaxia las estrellas son mucho más antiguas que el Sol. De haber vida en los planetas asociados a esas estrellas, puede que nos llevara mucha ventaja, con lo que correríamos un tremendo riesgo de ser aplastados por ella. Puede parecer una visión muy pesimista, pero no tenemos más que echar un vistazo a nuestro alrededor y ver lo que ha hecho el hombre con el resto de seres vivos de la Tierra. Lo normal es que los hayamos utilizado en nuestro provecho, así que ¿por qué deberíamos esperar un trato mejor hacia nosotros por parte de los posibles seres extraterrestres que se nos aproximaran? Lo cierto es que la vida de la Tierra “come o es comida” y no es esperable que en otros lugares la situación sea distinta.

Arthur C. Clarke decía: "Existen dos posibilidades: que estemos solos en el Universo o que no lo estemos. Ambas son igual de aterradoras". A pesar de eso, estamos seguros de que la misma inquietud que llevó a nuestros antepasados a salir de África y aventurarse en otros continentes nos llevará a nosotros a explorar el espacio, más y más lejos cada vez... Y así hasta que encontremos lo que hemos imaginado y soñado desde que fuimos conscientes de nuestra existencia, miramos hacia las estrellas y nos preguntamos qué podría haber en ellas.

Ester Lázaro Lázaro

Doctora en Ciencias Biológicas.

Investigadora Científica en el Centro de Astrobiología (CSIC-INTA).

Capítulo 18

La gravedad: ¿es una interacción fundamental?

¿Qué es una interacción fundamental?

Para empezar, unas palabras sobre interacciones en general. Hoy en día sabemos que la naturaleza, a nivel fundamental, es cuántica, por lo que ésta (la mecánica cuántica) es la explicación última de todas ellas. Sin embargo hablamos a veces de interacciones clásicas, y de interacciones de corto o de largo alcance. ¿Qué quiere decir esto? En el caso de la gravitación y el electromagnetismo, las interacciones son de largo alcance, y tienen manifestación clásica. Cuánticamente los efectos clásicos se entienden como efectos colectivos de una cantidad inconmensurable de cuantos. Las interacciones de corto alcance conocidas, como veremos en un momento, tienen un alcance tan pequeño que los efectos clásicos directos son imposibles de medir. Pero en principio podría haber interacciones de corto alcance que se manifestasen clásicamente⁴⁰ (por ejemplo, si el fotón tuviese una masa en un rango adecuado).

Volvamos a nuestro tema. Una interacción fundamental es aquella que no es reducible a otras más elementales. En la práctica y desde el punto de vista de la mecánica cuántica esto quiere decir que hay una partícula que es responsable de la interacción. Esta idea fue propuesta por primera vez por H. Yukawa⁴¹ en 1935. Curiosamente, su propuesta inicial consistía en explicar las fuerzas nucleares fuertes mediante el intercambio de piones. Hemos aprendido posteriormente que los piones no son partículas elementales, sino que están ellos mismos compuestos de un par quark-antiquark, y las fuerzas nucleares se deben al intercambio de gluones entre quarks.

Decimos que dos cuerpos interactúan mediante el intercambio de dicha partícula. Precisamente el alcance de la interacción es inversamente proporcional a la masa del cuanto intermediario. Largo alcance quiere decir masa nula (o más pequeña que la precisión de nuestras medidas). En este contexto hablamos de *fuentes* del campo en cuestión. O dicho de manera más precisa, estas partículas son los *quanta* del campo cuántico responsable de la interacción. Estos cuantos se pueden manifestar como partículas, que se producen y colisionan en los aceleradores como el CERN en Ginebra, o bien como campos clásicos cuando el número de cuantos es muy grande.

⁴⁰ Técnicamente, la palabra clásico, en este contexto, abarca efectos que no involucran directamente la constante de Planck.

⁴¹ Hideki Yukawa (1907-1981) fue un físico japonés que trabajó en la universidad de Kyoto, consiguiendo el premio Nobel en 1949.

Efectivamente, desde el punto de vista de la mecánica cuántica las ondas electromagnéticas están compuestas de un número inmenso de fotones. Todas las interacciones tienen esta doble naturaleza de onda/partícula, incluso en los casos en los que no existen campos clásicos por ser las interacciones correspondientes de corto alcance. Corto en este contexto quiere decir del orden del diámetro de un nucleón, 10^{-13} cm, una unidad de distancia llamada *fermi*. Cuando el cuanto es mucho más masivo (como es el caso de las interacciones débiles, que veremos en su momento), el alcance de la interacción es mucho más pequeño.

Naturalmente el concepto *fundamental* o *elemental*, como todos en ciencia, es provisional. Puede suceder que lo que hoy pensamos que es fundamental sea realmente el producto de una interacción más básica. Ya hemos visto que esto sucedió históricamente con los piones, que hoy se saben compuestos por una par quark-antiquark. La clave está en el famoso *principio de incertidumbre*. Para explorar distancias cada vez más pequeñas necesitamos energías más y más grandes. El valor de la energía máxima que se puede alcanzar en los aceleradores en un momento dado pone una cota sobre la escala de distancias más pequeña sobre la que podemos afirmar la no existencia de subestructura.

Las interacciones que se consideran fundamentales a las escalas que se han podido explorar experimentalmente son:

1) La *interacción electromagnética*. Es este otro ejemplo de lo que decíamos más arriba; durante siglos se pensó que la electricidad y el magnetismo eran dos interacciones fundamentales, hasta que Maxwell⁴² en 1865, basándose en los magníficos experimentos de Faraday, demostró que realmente eran la misma interacción. Modernamente se interpreta como debida al intercambio de *fotones*. Ya hemos visto como conjuntos macroscópicos de fotones constituyen las ondas electromagnéticas.

El descubrimiento de estas ondas ocupó buena parte del siglo XIX. Los primeros en detectarlas fueron los astrónomos: el infrarrojo (longitudes de onda más largas que el visible, y por consiguiente frecuencias más pequeñas; recordemos que la longitud de onda y la frecuencia son inversamente proporcionales) fue descubierto por el astrónomo William Herschel⁴³ ya en 1800, y el ultravioleta (la zona opuesta del espectro, a saber, longitudes de onda más pequeñas y por consiguiente, frecuencias más elevadas), por Wilhelm Ritter⁴⁴ en 1801. Una vez que Maxwell escribió sus ecuaciones, en 1887 Hertz⁴⁵ produjo artificialmente ondas de radio (longitudes de onda todavía más largas que el infrarrojo), y poco más tarde, Roentgen⁴⁶ produjo los rayos X, longitudes de onda más pequeñas que el ultravioleta.

La producción controlada del espectro completo de ondas electromagnéticas fue un paso previo al descubrimiento del fotón, postulado

⁴² James Maxwell (1831-1879) fue un físico escocés afincado en Cambridge, uno de los más relevantes del siglo.

⁴³ William Herschel (1738-1822) fue un astrónomo germano-británico, que llegó a ser astrónomo real.

⁴⁴ Wilhelm Ritter (1776-1810) fue un físico-químico alemán afincado en Jena.

⁴⁵ Heinrich Hertz (1857-1894) fue un físico alemán, director del Instituto de física de Bonn.

⁴⁶ Wilhelm Roentgen (1845-1923) fue un físico alemán afincado en la universidad de Würzburg.

por Einstein en su momento como una partícula de masa nula. Éste es un concepto absurdo desde el punto de vista de la física newtoniana, pero la relatividad especial explica la existencia de partículas con masa nula de manera consistente. Se mueven *siempre* a la velocidad de la luz, en cualquier sistema de referencia inercial. En física newtoniana, masa nula quiere decir momento nulo y energía nula. En la física relativista, en cambio, estas partículas tienen un momento lineal proporcional a su frecuencia (y por tanto, inversamente proporcional a su longitud de onda), y la constante de proporcionalidad es precisamente la constante de Planck. La energía de un fotón es proporcional a su momento, y por tanto, también a su frecuencia. Es decir, que tanto el momento como la masa son efectos cuánticos.

Experimentalmente fue el experimento de Compton⁴⁷ en 1923 el que convenció a la comunidad científica de que el fotón tenía un momento proporcional a su frecuencia, de acuerdo con las ideas de Einstein. Hoy en día la fórmula de Compton se interpreta como el resultado de la colisión de un fotón con un electrón en reposo. En esa colisión el momento del fotón cambia (por conservación de energía-momento), y por consiguiente también su frecuencia. Esto no es consistente con la interpretación de la luz como una onda clásica, en cuyo caso la frecuencia no cambiaría. Toda esta fascinante historia está muy bien resumida en el libro clásico de Pais [1].

2) Las *interacciones nucleares débiles*, responsables de las desintegraciones de los núcleos atómicos. Se interpretan como debidas al intercambio de los bosones vectoriales W^\pm y el Z^0 , descubiertos por Carlo Rubbia⁴⁸ y colaboradores en el experimento UA1 en el Laboratorio Europeo de Física de Partículas (CERN) en Ginebra en el año 1983. Para ello se hicieron chocar en los anillos de colisión del CERN haces de protones y antiprotones con una energía más de 300 veces la de un protón en reposo. A estas interacciones podemos añadir la mediada por el Higgs, que es una interacción cuántica de corto alcance, aunque es pronto para asegurar que el Higgs no tiene subestructura. Necesitamos explorar energías más grandes de las que somos capaces en la actualidad para acotar esta posibilidad.

3) Las *interacciones nucleares fuertes*. Éstas son las responsables de mantener a los protones y los neutrones unidos en los núcleos atómicos a pesar de la repulsión coulombiana debida al hecho de que los protones tienen todos carga positiva. Sucede que tanto los protones como los neutrones se interpretan como estados compuestos por tres quarks. Existe evidencia de que la interacción entre estos quarks está causada por el intercambio de 8 tipos de *gluones*, parecidos a los fotones, pero que interactúan entre sí. No existe indicación alguna de que los quarks tengan subestructura. Experimentalmente, la evidencia en favor de los quarks fue inicialmente obtenida en el acelerador SLAC, en Stanford, California, y la evidencia a favor de los gluones en PETRA, un experimento en el laboratorio DESY en Hamburgo.

4) Numerosos físicos piensan que la interacción gravitatoria podría ser interpretada en el mismo sentido que las otras interacciones, como causada

⁴⁷ Arthur Compton (1892-1962) fue un físico americano, profesor en la Universidad de Washington en St. Louis. Obtuvo el premio Nobel en 1927.

⁴⁸ Carlo Rubbia (1934-) es un físico italiano, que trabajó en la universidad de Harvard y en el CERN. Consiguió el premio Nobel en 1984.

por el intercambio de una partícula llamada en este caso el *gravitón*. La diferencia con los casos anteriores, es que de momento no hay evidencia directa de la existencia del gravitón, y de ahí las dudas sobre el carácter fundamental de la interacción gravitatoria.

En resumen: queremos ser lo más escépticos posible, y plantearnos la pregunta de si podemos al menos imaginar que no existen los gravitones, por lo cual, según nuestra definición anterior, la gravitación no sería después de todo, una interacción fundamental.

Primer Interludio: Los caminos a la Relatividad General.

El origen de la reflexión que llevo a Einstein⁴⁹ a elaborar en 1917 la Relatividad General es el famoso *principio de equivalencia*. Un campo gravitatorio es localmente equivalente a un sistema acelerado con una aceleración igual y opuesta a la de la gravedad en ese punto. O dicho de otra manera, en un sistema en caída libre el campo gravitatorio se anula. Aunque Einstein pensaba en términos de ascensores que se caían, hoy en día tenemos elocuentes manifestaciones de la veracidad de ese principio cuando vemos flotar en sus cápsulas a los astronautas en la estación espacial ISS (*International Space Station*).

Pero este camino no es el único posible. Mucho más tarde (en 1939) los físicos suizo-austríacos Fierz⁵⁰ y Pauli⁵¹ escribieron una ecuación para el gravitón consistente con la Relatividad Especial. Esta teoría es totalmente satisfactoria en tanto en cuanto no haya interacciones, pero presenta inconsistencias cuando se intentan incorporar éstas.

Esta línea de razonamiento, generalmente asociada al nombre de Feynman⁵², pero en la que participaron muchos autores incorporando pequeños detalles, conduce, después de resolver una complicada condición de autoconsistencia, exactamente a la misma teoría de la Relatividad General.

Curiosamente, en la vía de Feynman se parte de la física de gravitones; en la de Einstein del principio de equivalencia. Como vemos, la Relatividad General no necesita de, pero es consistente con, la existencia de gravitones.

¿Qué pasa con los gravitones?

En 2016 se comunicó a la comunidad científica la observación de ondas gravitatorias provenientes de la fusión de dos agujeros negros. Esta

⁴⁹ Albert Einstein (1879-1955) fue un importante físico alemán, que terminó afincado en Princeton.

⁵⁰ Markus Fierz (1912-2006) fue un físico suizo, asistente de Wolfgang Pauli.

⁵¹ Wolfgang Pauli (1900-1958) fue un físico austriaco, considerado como uno de los más creativos del siglo XX.

⁵² Richard Feynman (1918-1988) fue un físico americano, uno de los creadores de la electrodinámica cuántica.

observación, una de las más delicadas de la historia de la humanidad, fue posible gracias a la colaboración internacional LIGO. Tres de sus científicos (Weiss, Barish y Thorne)⁵³ fueron galardonados con el premio Nobel en 2017.

Es difícil exagerar la importancia de ese descubrimiento. Además de abrir una nueva ventana sobre el Universo, es claro que con el tiempo se aprenderá a modular y trabajar con este tipo de ondas, de modo análogo a lo que ocurrió con las ondas electromagnéticas.

Dada la pequeñez de la interacción gravitatoria comparada con las otras interacciones, parece sin embargo lejano el momento en que se pueda detectar el análogo del efecto Compton que demuestre fehacientemente la existencia de los gravitones. Exactamente ese mismo efecto es, en principio, observable, así como el correspondiente cambio de momento de los gravitones. Un cálculo sencillo sugiere sin embargo que estos efectos no pueden ser observados directamente hasta que las partículas tengan una energía 10^{19} (un 1 seguido de 19 ceros) veces de la de un protón en reposo.

Necesitamos nuevas ideas.

Segundo Interludio: Agujeros negros y materia oscura.

Antes de abordar el tema principal de este artículo, a saber, si es concebible que los gravitones no existan, tenemos que recordar unas ideas sobre entropía, agujeros negros y cosmología.

-La *entropía* de un sistema macroscópico es una medida del desorden. La entropía es en su origen un concepto termodinámico que tiene su explicación estadística, como la exponencial del número de estados cuánticos que corresponden a unas condiciones macroscópicas dadas (energía, etc.). Para sistemas en equilibrio termodinámico, la entropía es proporcional a la temperatura: cuanto más alta sea ésta, más desorden, y más entropía.

Los sistemas físicos al evolucionar tienden siempre a aumentar la entropía. Este principio general fue enunciado por primera vez por el físico austriaco Boltzmann⁵⁴ en el año 1872. Es intuitivo, y así ocurre para los sistemas ordinarios, que esa cantidad sea extensiva; esto es, proporcional al volumen del sistema. Esto quiere decir que cuanto más volumen, más estados cuánticos están disponibles para distribuir la energía del sistema.

-Veamos qué en un agujero negro, que no es ciertamente un sistema ordinario. Recordemos que un *agujero negro* es una región del espacio-tiempo

⁵³ Rainer Weiss (1932-) es un científico alemán que trabaja en el MIT en Cambridge (MA); Barry Barish (1936-) y Kip Thorne (1940-) son dos científicos americanos que actualmente trabajan en Caltech, en Pasadena (CA). Thorne es un teórico que contribuyó decisivamente al avance y popularización del proyecto. Barry Barish es originalmente un físico experimental de partículas elementales, que dinamizó notablemente el experimento LIGO desde que se incorporó a él en el año 1994. Lamentablemente el físico escocés Ron Drever (1931-2017), cuyas ideas fueron determinantes en la concepción del experimento LIGO, murió justo antes de la decisión del comité Nobel.

⁵⁴ Ludwig Boltzmann (1844-1906) fue un físico austriaco que trabajó en la universidad de Viena.

donde la atracción gravitatoria es tan intensa que nada, ni siquiera la luz puede escapar de él. Existe una superficie, llamada precisamente *horizonte* que separa los sucesos que pueden escapar de dicha atracción, y los que se ven ineludiblemente atrapados por el agujero negro y ya no pueden escapar.

Existe una analogía (que probablemente es más que una analogía, como veremos) termodinámica de un agujero negro. A todo agujero negro se le puede asociar una temperatura, la llamada *temperatura de Bekenstein-Hawking*⁵⁵, que es inversamente proporcional a la masa del agujero negro, y también una entropía, que es proporcional al área del horizonte de sucesos, que a su vez resulta ser proporcional al *cuadrado* de la masa de dicho agujero negro.

Se puede demostrar que cuando dos agujeros negros se fusionan en uno solo, el área del horizonte del agujero negro resultante es siempre mayor o igual que la suma de las áreas de los agujeros negros iniciales. Esta fue de hecho la primera pista de que la entropía del agujero era proporcional al área del horizonte.

Tenemos pues una entropía proporcional al área de la superficie que rodea al volumen que estamos considerando, en este caso el interior del agujero negro. Es conveniente subrayar que este es un fenómeno muy infrecuente. Ya hemos señalado que para los sistemas físicos ordinarios, a igualdad de todo lo demás, a mayor volumen, mayor entropía. Esto es lo que *no* es cierto para agujeros negros: la entropía es proporcional al área del correspondiente agujero, y no al volumen del interior del agujero negro.

Dado que esta dependencia del área es tan poco común, algunos físicos la han tomado como un principio fundamental de la gravedad a dichas escalas. Esta es precisamente la base o punto de partida del *principio holográfico* de Gerardus 't Hooft⁵⁶ y Leonard Susskind⁵⁷, que postula que los grados de libertad fundamentales a nivel cuántico han de distribuirse en una superficie o pantalla, como si se tratase de un holograma, en vez de distribuirse por todo el volumen del sistema.

La temperatura de Hawking por otra parte, es la temperatura de la radiación que emite un agujero negro debido a fluctuaciones cuánticas. En cambio la entropía no es la que correspondería a un *cuerpo negro* (que sería extensiva), como ya se ha mencionado, aunque se puede explicar en teoría de cuerdas para cierto tipo de agujeros cargados llamados *extremales*, que son precisamente aquellos que se pueden embeber en una teoría supersimétrica. Explicar, en este sentido, quiere decir que se han identificado los estados cuánticos precisos que dan lugar a dicha entropía. Ya decíamos antes que toda la termodinámica de agujeros negros puede ser más que una analogía. Pero todavía está por demostrar que esto es cierto para agujeros negros no supersimétricos, que son los únicos de interés astrofísico.

En el espacio-tiempo puede haber horizontes que no son debidos a la presencia de un agujero negro. Por ejemplo, un observador uniformemente acelerado no es capaz de interactuar con todo el espacio-tiempo; se dice que tiene un horizonte (en este caso, llamado de Rindler). Esto quiere decir dos

⁵⁵ Jacob Bekenstein (1947-2015) fue un físico mexicano-israelí que trabajó en la universidad hebrea de Jerusalén. Stephen Hawking (1942-2018) fue un físico británico, asociado a la universidad de Cambridge (UK).

⁵⁶ Gerardus 't Hooft (1946-) es un físico holandés que trabaja en la universidad de Utrecht y consiguó el premio Nobel en 1999.

⁵⁷ Leonard Susskind (1940-) es un físico americano que trabaja en la universidad de Stanford.

cosas. Por una parte, hay un trozo del espacio-tiempo que *nunca* podrá interactuar con el observador acelerado (pero éste sí que puede mandar señales a esa región). Por otra, existe una región diferente, a la que el observador acelerado nunca podrá mandar ninguna señal (pero el observador sí que puede recibir señales provenientes de esta segunda región). Hay además una tercera región en la que ni se pueden recibir ni tampoco emitir señales hacia el observador acelerado.

Pues bien, siempre que hay un horizonte, hay una temperatura asociada, así como todas las cantidades termodinámicas que se deducen de ella, en particular la entropía. La temperatura del horizonte es proporcional a la aceleración del sistema acelerado que es localmente equivalente al campo gravitatorio, según el principio de equivalencia que acabamos de recordar en el primer interludio.

-En los años 1960 Vera Rubin y Kent Ford⁵⁸, basándose en trabajos anteriores, demostraron concluyentemente que la masa visible no era suficiente para mantener las estrellas de la galaxia ligadas gravitacionalmente a ella. Aceptando estos datos, hay dos explicaciones. O bien hay más masa de la que se ve, y a esta hipotética masa le llamaríamos *masa oscura*; o bien las leyes de Newton dejan de ser válidas a las enormes escalas características de las galaxias.

Es un tema recurrente desde los trabajos del físico israelí Milgrom (1983) que una modificación *ad hoc* de la dinámica newtoniana podría ser una explicación alternativa a la materia oscura para solucionar el problema de las curvas de rotación en galaxias. Con las observaciones recientes de microlentes gravitatorias, se hace cada vez más complicado construir modificaciones de la gravedad consistentes con todos los datos; pero por otra parte es cierto que no se ha encontrado (¿todavía?) en la naturaleza un candidato viable para la materia oscura.

El decidir cuál de estas dos explicaciones posibles es correcta solo será posible mediante observaciones más precisas, como por ejemplo la detección directa de dicha materia oscura o bien la verificación de alguna predicción que nos obligue a modificar la Relatividad General.

¿Puede ser la gravedad una fuerza entrópica?

¿Qué es una fuerza entrópica? Es éste un concepto que se usa en general para sistemas muy complicados (por ejemplo los polímeros) de los cuales no se tiene una comprensión detallada a partir de primeros principios, pero de los que sabemos, de acuerdo con la segunda ley de la termodinámica, que su entropía aumenta. A las fuerzas que pueden interpretarse como causantes de ese aumento se les llama precisamente *entrópicas*.

La idea básica del físico americano Ted Jacobson⁵⁹ en el año 1995 (retomada con ligeras variaciones, por el físico indio, en el año 2010, Thanu

⁵⁸ Vera Rubin (1928-2016) fue una astrónoma americana asociada sobre todo a la *Carnegie Institution*. Kent Ford (1931-) es un astrónomo americano.

⁵⁹ Theodore Jacobson (1954-) es un físico americano que trabaja en la Universidad de Maryland [2].

Padmanabhan⁶⁰ y el holandés Erik Verlinde⁶¹ en el año 2011) es aplicar el principio de que en todo sistema en el que existe una entropía no uniforme, existe una *fuerza entrópica*, proporcional a la variación espacial de la entropía; aplicar esa idea, decimos, al campo gravitatorio, utilizando para ello las nociones de entropía y de temperatura que acabamos de explicar brevemente.

El punto de vista de los autores mencionados (especialmente de Verlinde) se puede glosar de la siguiente manera: no nos pronunciamos sobre los grados de libertad cuánticos fundamentales, sino que intentamos sacar consecuencias generales sobre identidades termodinámicas en el horizonte de un agujero negro.

En resumen, en mi opinión, lo que estos interesantes trabajos demuestran es que *sea cual sea la teoría subyacente*, las ecuaciones de Newton son más o menos inevitables a nivel macroscópico. De ahí a las ecuaciones de Einstein se puede llegar postulando covarianza relativista, siguiendo el camino de Fierz-Pauli-Feynman (recordemos que no fue históricamente el seguido por Einstein).

En este sentido, se puede decir que la Relatividad General es análoga a una *teoría emergente* de las que utilizan los físicos para describir la materia condensada (como los plasmones, por ejemplo). Interacciones emergentes son las que aparecen cuando no se está interesado (o no se puede tener en cuenta debido al número muy elevado de grados de libertad) en el detalle de las interacciones básicas, y se promedia sobre ciertas cantidades o regiones del sistema. Las teorías emergentes más conocidas aparecen cuando se estudian objetos macroscópicos, como en materia condensada, o en física estadística.

El uso de este concepto en física teórica es una interesante novedad.

Este tipo de análisis tiene dificultades⁶² para considerar sistemas dependientes del tiempo, como los que son inevitables cuando se trata de describir todo el universo, es decir, el problema Cosmológico. De hecho, en un segundo trabajo, Verlinde postula que hay una contribución extensiva (proporcional al volumen) a la entropía que es responsable de modificaciones de la gravedad que podrían hacer innecesaria la materia oscura.

En definitiva, nuestro punto de vista es que la búsqueda de las variables fundamentales de la naturaleza a nivel cuántico sigue adelante. Lo que estos trabajos *revisionistas* señalan es la existencia de propiedades independientes de los detalles de estas variables.

De hecho algunos autores, como por ejemplo el físico chino Tower Wang entre otros han propuesto una derivación entrópica de las ecuaciones de Maxwell⁶³. Esto demuestra al menos que el hecho de que exista una manera

⁶⁰ Thanu Padmanabhan (1957-) es un físico indio que trabaja en el IUCAA, en Pune [3].

⁶¹ Erik Verlinde (1962-) es un físico holandés asociado a la universidad de Amsterdam [4].

⁶² Recomendamos la lectura del análisis de Carlip [5] para profundizar en las dificultades de la idea general.

⁶³ Incidentalmente este artículo concreto [6] apareció publicado *antes* que el trabajo de Verlinde.

entrópica de entender una interacción no impide necesariamente la existencia de quanta del campo correspondiente.

Está claro que para entender completamente la gravitación como debida al intercambio *cuántico* de gravitones, necesitamos una teoría que contenga a la vez a la gravitación y a la cuántica. La búsqueda de dicha teoría que haga compatibles (o que posea como ciertos límites) la Relatividad General y la Mecánica Cuántica es una de las aventuras intelectuales más fascinantes que se pueden emprender.

Es posible que existan continentes enteros de conocimiento que están esperando a ser descubiertos. Mientras tanto, el viaje continúa en medio de una densa niebla.

Referencias:

- [1] Abraham Pais, "Inward bound: Of matter and forces in the physical world." (Clarendon Press, Oxford, 1988)
- [2] T. Jacobson, "Thermodynamics of space-time: The Einstein equation of state," Phys. Rev. Lett. **75** (1995) 1260 doi:10.1103/PhysRevLett.75.1260 [gr-qc/9504004].
- [3] T. Padmanabhan, "Thermodynamical Aspects of Gravity: New insights," Rept. Prog. Phys. **73**, 046901 (2010)
- [4] E. P. Verlinde, "On the Origin of Gravity and the Laws of Newton," JHEP **1104**, 029 (2011) doi:10.1007/JHEP04(2011)029 [arXiv:1001.0785 [hep-th]]. "Emergent Gravity and the Dark Universe," SciPost Phys. **2** (2017) no.3, 016 doi:10.21468/SciPostPhys.2.3.016 [arXiv:1611.02269 [hep-th]].
- [5] S. Carlip, "Challenges for Emergent Gravity," Stud. Hist. Phil. Sci. B **46**, 200 (2014) doi:10.1016/j.shpsb.2012.11.002 [arXiv:1207.2504 [gr-qc]].
- [6] T. Wang, "The Coulomb Force as an Entropic Force," Phys. Rev. D **81**, 104045 (2010) doi:10.1103/PhysRevD.81.104045 [arXiv:1001.4965 [hep-th]].
- [7] E. Alvarez, "Quantum Gravity: An Introduction To Some Recent Results," Rev. Mod. Phys. **61**, 561 (1989) doi: 10.1103/Rev.Mod.Phys.61561

Enrique Álvarez

Catedrático de la UAM y miembro permanente del IFT.
Departamento de Física Teórica, Instituto de Física Teórica,
Universidad Autónoma de Madrid, IFT-UAM/CSIC.

Capítulo 19

El LHC: acelerador de partículas, acelerador del conocimiento.

Durante el periodo de aproximadamente 40 años que va desde 1973, con el descubrimiento en el CERN de las corrientes débiles neutras, hasta 2012, en que se descubre, también en el CERN, el bosón de Higgs, tuvo lugar una serie de observaciones experimentales en distintos aceleradores de partículas, que provocaron una fuerte expansión de nuestro conocimiento de la física. Fue el periodo de consolidación de lo que se conoce con el nombre de teoría de unificación electrodébil, que forma la parte más característica del llamado (de forma un tanto displicente) Modelo Estándar de la física de partículas.

Vamos a resumir cuáles fueron las ideas, e hitos principales, de este desarrollo previo, que sin duda estuvo en el origen del actual colisionador LHC (Large Hadron Collider) del CERN. Si hubiera que asociar un nombre a la parte teórica de este desarrollo (cosa a veces necesaria, para fijar nuestras ideas), yo lo asociaría con Steven Weinberg, autor del principal artículo seminal para la unificación electrodébil en 1967 [1].

Hasta entonces se conocía la interacción débil, que gobierna la desintegración beta de los núcleos radiactivos emisores de electrones (transmutando un protón en neutrón, o viceversa), y también la desintegración de muones (que son electrones pesados) en electrones. En ambos casos se produce la emisión de neutrinos. La interacción débil es apenas conocida por el gran público, que tiende a confundirla con la energía nuclear. Esta confusión es comprensible, porque ambas operan en el entorno de los núcleos. La primera forma parte inevitable de los ciclos de reacciones termonucleares que provocan, por ejemplo, la explosión de las supernovas, o la combustión del hidrógeno solar para formar Helio. No hace falta decir lo bien que se conocía en 1967 el electromagnetismo, que gobierna, entre otras cosas, las ondas de radio y televisión, así como también la energía química de los átomos y de las moléculas. Mientras el electromagnetismo transmite su energía por partículas neutras (los fotones), la interacción débil lo hace por partículas cargadas (el bosón W emitido por el protón o neutrón, que es luego convertido en electrón y neutrino) [2] [3].

El primer paso dado por Weinberg fue darse cuenta de que la interacción débil tenía que admitir también una versión eléctricamente neutra, donde los neutrinos pudiesen golpear electrones y núcleos, sin necesidad de intercambiar con ellos bosones W cargados. La partícula intercambiada sería un bosón eléctricamente neutro, como el fotón (llamado W^3). De esta forma tendríamos una terna de bosones (W^+ , W^- y W^3) que otorgarían a la interacción débil una simetría similar al electromagnetismo, pero con 3 fotones en lugar de uno solo.

Toda la belleza del electromagnetismo, condensada en las famosas ecuaciones de Maxwell, podía ahora extenderse a la interacción débil, simplemente admitiendo que, en total, necesitamos tener 4 campos similares al campo electromagnético (W^+ , W^- , W^3 y B), en lugar de únicamente el fotón. Tres de ellos para dar cuenta de la interacción débil, y otro más B , de cuyo acoplo surgiría el electromagnetismo. Weinberg se dio cuenta de que no era posible la simple yuxtaposición de ambos tipos de interacción, sino que los campos neutros W^3 y B debían estar mezclados, de tal manera que el fotón A y la corriente débil neutra observada Z^0 fuesen combinaciones lineales ortogonales de ambos.

Viendo que la simetría a cuatro así obtenida era una extensión de la ya conocida para el electromagnetismo, y que los cálculos funcionaban, Weinberg se enfrentó al verdadero problema de fondo, el más difícil de resolver: mientras la interacción débil actuaba solo a distancias muy cortas (por debajo del femtómetro), en cambio el electromagnetismo actuaba a distancias tan largas como se quiera (piénsese en las ondas de radio). ¿Cómo podría admitirse que ambos tuviesen igual naturaleza? La forma de describir esto en la teoría cuántica y relativista es constatar que mientras los bosones de la interacción débil tienen masa muy elevada, en cambio el fotón tiene masa nula. Y la dificultad estriba en el hecho de que unos tengan masa y otros no, pues ello rompe la simetría electrodébil.



Fig.1. Steven Weinberg, Premio Nobel 1979, Archivo Fundación Nobel.

La solución la encontró Weinberg utilizando adecuadamente un principio ya establecido por entonces en la teoría cuántica de los campos, que era el de atribuir al vacío (el estado sin partículas que lo llena todo) la responsabilidad de una simetría rota, como idea genérica. Él lo aplicó a este caso concreto, postulando la existencia de 4 nuevos campos sin espín, también dotados de carga débil, que interactúan con bosones, electrones y quarks dentro de la simetría electrodébil. Uno de ellos (el llamado bosón de Higgs) toma un elevado valor en el vacío. Es decir, que el vacío no es inerte, sino que contiene, distribuida por todo el espacio, carga débil neutra, y hace así posible que la interacción débil tenga corto alcance, sin alterar la simetría de las interacciones. Esto se conoce en la literatura como mecanismo de Higgs. Además, Weinberg explica brillantemente porqué el fotón se queda sin masa, pues la carga eléctrica que comunica no está, por el contrario, presente en el vacío. Este hecho da origen a la singularidad del electromagnetismo. El otro bosón neutro que sí adquiere masa es el Z^0 . Como hemos dicho, ambos son rotaciones ortogonales de los campos neutros W^3 y B , habiendo tomado el ángulo de rotación θ_W el nombre de su descubridor.

Weinberg se adelantó bastante a su tiempo, pues fue necesario ese periodo de 40 años que mencionaba antes, para que varias generaciones de aceleradores, en distintas partes del mundo, fuesen desvelando los aspectos esenciales de la teoría. En 1984 se descubrieron los bosones W^+ y Z^0 en el CERN. En este proceso se fueron descubriendo, además, nuevos quarks, incluyendo los más pesados: el charm (c) en 1974, el bottom (b) en 1977 y el top (t) en 1992. Es curioso que, siendo su teoría universal para quarks y electrones, Weinberg no mencionase para nada los quarks en su modelo inicial. Esto revela hasta qué punto el avance de la teoría se debió a un esfuerzo conjunto de muchos grandes físicos.

$$\begin{array}{ccc}
 \begin{pmatrix} u \\ d \end{pmatrix} \begin{array}{l} 2 \text{ MeV}/c^2 \\ 4 \text{ MeV}/c^2 \end{array} &
 \begin{pmatrix} c \\ s \end{pmatrix} \begin{array}{l} 1300 \text{ MeV}/c^2 \\ 10 \text{ MeV}/c^2 \end{array} &
 \begin{pmatrix} t \\ b \end{pmatrix} \begin{array}{l} 173000 \text{ MeV}/c^2 \\ 5100 \text{ MeV}/c^2 \end{array}
 \end{array}$$

Fig.2. Las tres familias conocidas de quarks y sus masas.

Hablando de los quarks, debemos mencionar aquí otro avance esencial de la teoría electrodébil, cuya contribución mayor, en este caso, fue debida a los físicos japoneses Kobayashi y Maskawa en 1973[4]. Se trataba de explicar, en dicha teoría, porqué la tasa de desintegración de los quarks con su espín a derechas es distinta de la de los antiquarks con su espín a izquierdas. Este fenómeno se llama en la literatura violación de la simetría CP (carga-paridad), o simplemente violación de la simetría materia/antimateria. No siendo incompatible con la relatividad restringida de Einstein, digamos que resulta un tanto extraño a ella, prueba de ello es que en la teoría electrodébil no se encontró cabida fácil para él. Kobayashi y Maskawa dieron una explicación vanguardista del fenómeno en base a la existencia de tres réplicas distintas de quarks, y al hecho de que estos quarks tuviesen masas distintas. En esto se adelantaron también a su tiempo, pues por aquél entonces no se había descubierto todavía ninguna de las generaciones de quarks pesados.

Resulta que esa ruptura de la simetría materia/antimateria es absolutamente clave para explicar porqué las galaxias de nuestro universo están todas compuestas por materia, y ninguna por antimateria, como parece ser el caso. En los primeros instantes del universo, a muy alta temperatura, es natural pensar que hubo igual número de quarks y antiquarks. Un lapso de tiempo más tarde estos se aniquilaron entre sí, dando lugar a enormes cantidades de energía. Esto hubiera hecho desaparecer los protones y neutrones que forman las galaxias, si no fuese por el hecho de que, previamente, se produjo un pequeño exceso de materia sobre antimateria. Este exceso, de aproximadamente 10^{-9} , se denomina *asimetría bariónica cosmológica*, y requiere inevitablemente de la violación de la simetría CP. Sin embargo, la teoría de Kobayashi-Maskawa, aplicada a muy altas temperaturas, no permite explicarlo, pues esta produce una violación CP mucho menor. Por tanto, tenemos aquí un misterio que resolver, y muchos físicos pensábamos, al inicio del LHC, que podríamos resolver este misterio encontrando nuevas formas de violación de la simetría CP, que fuesen más allá de dicha teoría, utilizando para ello quarks pesados.

Durante las décadas de 1980 y 1990 se construyeron dos aceleradores consecutivos que parecían ser las máquinas idóneas para dilucidar todos los aspectos anteriores de la unificación electrodébil. En ambos casos se trataba de anillos de colisión para producir la aniquilación electrón-positrón, a energías muy elevadas, y disponiendo de corrientes también muy elevadas de antimateria (los positrones), que llegaron a alcanzar decenas de miliamperios. El primero, PETRA, operó en el laboratorio de DESY (Hamburgo) en el periodo 1979-1984, y el segundo, LEP, en el CERN, durante el periodo 1989-2001. De hecho, el segundo fue prácticamente una copia del primero, duplicando su energía y trasladando al CERN su centro de operaciones. Hay que decir que el túnel de 27 Km que hubo que excavar en la región próxima a Ginebra para alojar LEP es el mismo que hoy se utiliza para el LHC, y no se trata de una obra civil menor.

Ambos aceleradores tuvieron enorme éxito, pues pudieron establecer con gran precisión la mayor parte de los aspectos cruciales de la teoría de unificación electrodébil, incluyendo una auténtica radiografía del bosón Z^0 , y del famoso ángulo de Weinberg θ_W de mezcla entre los bosones W^3 y B . Todo encajaba a la perfección, pero faltaba una pieza esencial: el bosón de Higgs no aparecía por ninguna parte. Hay que decir también que, además de dejar casi establecida la teoría de unificación electrodébil, estos aceleradores realizaron nuevos descubrimientos de primera fila: el del gluón, en PETRA, como partícula real que media la interacción fuerte, y el de la existencia de solo tres réplicas de neutrinos ligeros, en LEP.

Así pues, durante la década de 1990 se hizo acuciante la idea de construir un nuevo gran acelerador de partículas, cuyo objetivo primario fuese dilucidar la existencia del bosón de Higgs. De hecho, ya en el laboratorio Fermilab (en Chicago) se puso en operación un acelerador protón-antiprotón con este fin, el Tevatrón, pero se quería tener una máquina aún más potente, para garantizar el éxito del proyecto. La idea que finalmente se impuso en la comunidad internacional, provino del CERN: reutilizar el túnel de 27 Km de LEP para instalar en él un nuevo acelerador protón-protón, a una energía de colisión muy elevada (16 TeV).

A diferencia del caso del bosón W , cuya masa pudo ser inferida con antelación a partir del corto alcance de la interacción débil (la constante de Fermi G_F), la teoría electrodébil no es capaz de predecir la masa del bosón de Higgs, y por eso era necesario dotar al acelerador de una energía suficientemente elevada, que no dejase dudas sobre la existencia o no del bosón. Utilizar exclusivamente protones, en este sentido, en lugar de provocar la aniquilación protón-antiprotón (como en el Tevatrón), se juzgó conveniente para poder elevar todavía más la intensidad de los haces colisionantes. La decisión de construir la máquina en el CERN, tomada en 2001 por su entonces director general, el físico italiano Luciano Maiani, no estuvo exenta de polémica. Muchos pensaron en aquel momento que era más útil proseguir la toma de datos con LEP, aumentando progresivamente su energía y su intensidad. Visto en retrospectiva, hoy podemos decir que la decisión tomada fue un éxito rotundo.

Hubo otros aspectos que influyeron también en la construcción del acelerador. Las observaciones en astrofísica por entonces ya ponían de manifiesto con claridad la necesidad de introducir la llamada materia oscura, tanto para describir la velocidad de rotación de las galaxias, como para dar cuenta del fenómeno de las imágenes dobles originadas por la gravitación de Einstein. No es evidente en absoluto que ninguna de las partículas que forman parte de la teoría de unificación electrodébil pueda dar cuenta de la materia oscura. Las observaciones en astrofísica son consistentes con partículas de masa elevada (del orden de los núcleos atómicos) que interactúan muy débilmente. Sin embargo, en teorías que van más allá del Modelo Estándar, sí pueden encontrarse partículas con estas características. Son tremendamente populares en el campo de la física teórica las teorías llamadas supersimétricas, que postulan la existencia de réplicas de todas las partículas elementales en otras, que tienen el valor opuesto de su espín. Opuesto quiere decir aquí que cambiamos su valor entero por otro semientero, y viceversa. Es decir, que transformamos bosones en fermiones, y recíprocamente, en esto consiste la supersimetría.

Pues bien, en el campo hipotético de la supersimetría sí existen partículas que pueden jugar el papel de la materia oscura, como por ejemplo los neutralinos. La posibilidad de producir estas partículas en el acelerador, con energía y luminosidad adecuadas, se contempló, y se contempla actualmente, como un factor determinante para el desarrollo del proyecto. Bien es verdad que, de ser cierta la supersimetría, no solamente los neutralinos se manifestarían en el laboratorio, sino otras muchas nuevas partículas en la región de masa del TeV/c^2 que otorgan al acelerador un potencial de descubrimiento espectacular.

Téngase bien en cuenta que, desde el punto de vista teórico, el actual Modelo Estándar presenta deficiencias notables, en lo que respecta, sobre todo, a la llamada física del sabor. Es decir, a la existencia de distintas generaciones de quarks y de leptones (análogos al electrón), con masas muy distintas. La jerarquía de masas de las partículas elementales no es comprendida en absoluto por el Modelo Estándar.

Podemos, entonces, resumir en 3 las principales preguntas que motivaron la construcción del LHC, relativas a la unificación de las fuerzas fundamentales:

- ¿Es realmente el vacío el responsable de la ruptura de la simetría electrodébil, a través del bosón de Higgs? O, por el contrario, debemos pensar en otras alternativas, como, por ejemplo, nuevas dimensiones espaciotemporales, en el dominio de la gravitación.
- ¿Qué es la materia oscura? ¿Puede ser asimilada a partículas pesadas que se produzcan en el acelerador, más allá del Modelo Estándar? ¿Es cierta la supersimetría?
- ¿Porqué el universo adquirió más materia que antimateria en los primeros instantes, y no existen galaxias de antimateria? ¿Existen nuevas fuentes de violación de la simetría CP que el acelerador pueda poner de manifiesto?

A estos objetivos es necesario añadir otro más, de no menor importancia, como es el mejor conocimiento que puede aportar el LHC de la interacción fuerte (o nuclear), descrita por las interacciones de quarks y gluones. En particular, dado que la máquina puede también acelerar *iones*, pesados o ligeros, se abre la posibilidad de estudiar estados de la materia nuclear con mayor número de quarks pero, sobre todo, a mucha mayor temperatura T (por la elevada energía de las colisiones), que en proyectos anteriores. Por esta razón se marcó como objetivo adicional estudiar las propiedades del llamado plasma de quarks y gluones, observado anteriormente en el acelerador americano RHIC en el BNL (New York), y en el propio CERN. Esta línea justificó desde el principio un experimento dedicado, ALICE, que hace uso del acelerador en periodos de tiempo cortos, pero suficientes para sus fines. Téngase en cuenta que las colisiones tienen en este caso una probabilidad de interacción mucho mayor, haciendo innecesarios haces muy intensos.

La tecnología del acelerador fue, hasta cierto punto, una extrapolación de las ya conocidas de aceleradores anteriores. Tanto los imanes dipolares, necesarios para curvar los protones sobre la circunferencia de 27 Km, como los cuadrupolos y sextupolos, necesarios para focalizarlos, así como las cavidades de radiofrecuencia que los aceleran, fueron diseñados para un bajo consumo energético.



Fig.3. Imanes dipolares del LHC descendiendo a 100 m de profundidad para su instalación. Fuente: CERN.

Esto exigió utilizar materiales superconductores, adoptándose la temperatura de operación del Helio líquido en todos los elementos del acelerador. Por su elevado número, esto supone una utilización masiva de criogenia, plantas de refrigeración para producir Helio líquido, y cables de conexión superconductores, que supera a cualquier otra aplicación industrial.

Se facilitaron cuatro áreas experimentales para las colisiones, que utilizaban, de hecho, la obra civil de las ya existentes para el acelerador de LEP. Dos de ellas se dedicaron a experimentos de propósito general, concurrentes en sus objetivos de descubrir el Higgs y las partículas supersimétricas, pero diferentes en su diseño experimental: ATLAS y CMS. La

tercera se dedicó a un experimento focalizado en aprovechar el elevadísimo flujo de quarks b a bajo ángulo, para estudiar con mayor precisión las asimetrías materia/antimateria y la física del sabor: LHCb. La cuarta se dedicó a analizar las colisiones de iones pesados, y el plasma de quarks y gluones: ALICE.

Tanto ATLAS como CMS fueron diseñados para aceptar los productos de las colisiones protón-protón sobre todo el ángulo sólido 4π , y utilizar la máxima intensidad de los haces de protones, aún produciéndose múltiples interacciones de interés en cada entrecruzamiento de los pulsos. Para poder

descubrir el Higgs y las nuevas partículas pesadas, era necesario reconstruir su masa, y ello requiere ser sensible a todas las direcciones de emisión posibles de sus productos de desintegración. Téngase en cuenta que la probabilidad de formación de partículas muy pesadas disminuye dramáticamente al aumentar su masa. Además, la utilización del bosón W como sonda de prueba exige reconstruir con precisión la energía no detectada de sus neutrinos, lo cual requiere detectores herméticos, que no dejen huecos en ninguna dirección. Todo esto encarece notablemente la construcción, si se tiene en cuenta además que los detectores deben ser resistentes a la radiación, dado el importante flujo de neutrones lentos que se produce en las zonas más internas. Sin llegar a compararse con los de una central nuclear, son exigibles medidas de protección radiológica sobre los materiales irradiados.

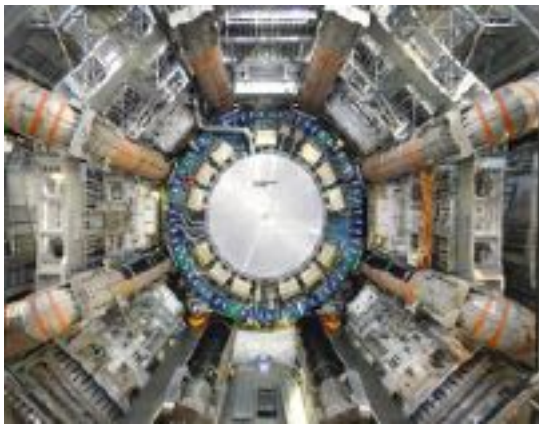


Fig.4. La parte central del experimento ATLAS, con el calorímetro de Ar líquido. Fuente: CERN.



Fig.5 El solenoide del experimento CMS, rodeado de cámaras de muones entre el hierro magnetizado. Fuente: CERN.

Dentro de las condiciones generales anteriores, ATLAS y CMS adoptaron diseños bien distintos para sus detectores. Mientras ATLAS abordó la construcción de ambiciosos imanes toroidales superconductores en sus zonas externas, CMS se limitó a un dipolo solenoidal, también superconductor, con entrehierro magnetizado que permite el seguimiento de los muones (electrones pesados penetrantes) en su interior. Nótese que la medida magnética del momento relativo entre dos partículas es esencial para la determinación precisa de su masa. Ambos experimentos dedicaron especial atención a otro elemento altamente no trivial: la detección de pares de fotones opuestos, modo de desintegración esperado del Higgs, que requiere alta resolución en la masa. Mientras CMS adoptó elevadas cantidades de Plomo-Wolframio, ATLAS construyó un calorímetro de Argón líquido, detectores novedosos ambos, con sofisticados sistemas de extracción de la luz de centelleo.

El caso de LHCb es bien distinto a los anteriores. Al tratarse de un detector de bajo ángulo, la detección de la energía faltante de los neutrinos se hace muy difícil, y no tiene mucho sentido que el detector sea hermético. Por el contrario, el énfasis se puso en un detector de microvértice de gran profundidad (22 planos), que aprovecha el mayor vuelo que tienen los quarks pesados en la dirección hacia delante, por la dilatación temporal relativista (los quarks b

vuelan típicamente varios mm antes de desintegrarse, con una vida media en torno al picosegundo). Se maximizó la resolución en la masa con un imán dipolar vertical convencional, así como el recubrimiento para muones (partículas muy penetrantes) de bajo momento transverso respecto al haz. Además se instalaron detectores de luz Cherenkov (RICH) que permiten observar el anillo de fotones asociado a cada partícula, y distinguir muy bien piones de kaones (es decir, la presencia del quark s en las desintegraciones). Estos detectores no podrían instalarse fácilmente en una geometría con recubrimiento 4π . La detección de vértices secundarios (enfocada a los quarks b y c) se convierte en elemento distintivo, con la contrapartida de no poder aceptar múltiples interacciones en cada colisión. Esto obliga a tomar datos con una intensidad del haz más reducida, lo cual se consigue modificando la óptica del acelerador en esta área experimental, para que la toma de datos pueda ser simultánea con ATLAS y CMS.



Fig.6. El imán convencional del experimento LHCb, que permite separar materia de anti-materia. Fuente: CERN.



Fig.7. El experimento ALICE estudia las colisiones ión-ión. Los detectores se encuentran dentro de un solenoide horizontal. Fuente: CERN.

El inicio de las colisiones, que tuvo lugar en 2009, resultó curiosamente fallido. Las conexiones superconductoras en determinados imanes no pudieron soportar la muy elevada energía prevista de 16 TeV, volviéndose conductores normales. Esto provocó algunas explosiones por la elevación súbita de temperatura, un accidente perfectamente previsto en los protocolos de seguridad de las alimentaciones eléctricas. Fue necesario, eso sí, un rediseño de las conexiones, y en 2010 se reiniciaron sin problemas las colisiones, aunque solo a 7 TeV, y con una intensidad reducida.

Las tomas de datos en 2011 y 2012 con una energía ya aumentada a 8 TeV, y con la intensidad prevista, produjeron el primer gran logro: el descubrimiento del bosón de Higgs con una masa de $125 \text{ GeV}/c^2$. El anuncio a la comunidad internacional tuvo lugar el 4 de Julio de 2012 por parte de las colaboraciones ATLAS y CMS. El modo de desintegración observado inicialmente fue precisamente en dos fotones, el más difícil, como ya hemos comentado. Poco después se observó también su desintegración, con la misma masa, en pares de bosones W^+W^- y Z^0Z^0 . Es curioso que puedan observarse estas desintegraciones, a pesar de que la suma de masas de los bosones sea

superior a $125 \text{ GeV}/c^2$, pero ello es posible gracias a las fluctuaciones cuánticas.

Tras una parada del acelerador en el periodo 2013-2014, se reanudó la toma de datos en 2015-2018 (Run 2), alcanzándose la energía de 14 TeV, muy cercana a la originalmente prevista. La operación del acelerador y de los detectores ha sido excelente durante este periodo, llegando a superarse los valores previstos de luminosidad integrada en todos los experimentos. Ello ha permitido realizar cientos de publicaciones científicas que no solamente han extendido enormemente nuestro conocimiento de la física de partículas, sino que han cumplido una buena parte de los objetivos iniciales, explorando distancias muy cortas y poniendo a prueba la estructura interna de quarks, leptones y bosones. Y también, como en máquinas anteriores, se han producido algunos nuevos descubrimientos, fuera del programa inicial.

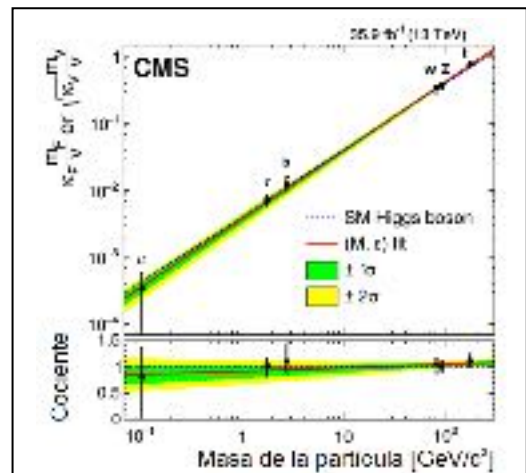


Fig.8. Constantes de acoplamiento del bosón de Higgs a cada una de las partículas pesadas en que se ha medido, y predicción de la teoría de unificación electrodébil. Se indica abajo el cociente con la predicción.
Fuente: Colaboración CMS, arXiv:1809.10733 [hep-ex] (2018).

En Diciembre de 2018 acaba de iniciarse la segunda parada larga del acelerador, que durará hasta 2021, y que permitirá mejorar de manera importante su operación futura, incluyendo todos los detectores. Es un momento adecuado para realizar un primer balance de resultados positivos para el LHC en su conjunto, que podemos resumir como sigue:

- 1) A los modos de desintegración inicialmente observados del bosón de Higgs H^0 en fotones y bosones W/Z^0 , se han añadido, de forma crucial, sus desintegraciones en quarks y electrones pesados (leptones). Han quedado establecidos tres modos de desintegración en pares partícula-antipartícula, por ATLAS y CMS en 2018: quarks b, leptones τ (el electrón más pesado), y quarks t (el quark más pesado). El acoplamiento de este último ha podido ser determinado al observarse la producción asociada $H^0 t$. Los tres acoplos anteriores han resultado estar en buen acuerdo con la predicción de la teoría electrodébil, que señala un acoplamiento proporcional a sus masas. Esto viene a confirmar que la partícula descubierta se ajusta a su papel previsto en la ruptura de la simetría electrodébil, de otorgar masa no solo a los bosones, sino también a quarks y leptones[5][6], como mostramos en la Fig. 8.
- 2) El descubrimiento de que no hay partículas supersimétricas con masas por debajo de aproximadamente $1 \text{ TeV}/c^2$, dependiendo del tipo de partícula y, en algunos casos, del modelo de supersimetría. Se han publicado múltiples búsquedas de este tipo de partículas, por los experimentos ATLAS y CMS, todas ellas con resultado negativo. Las búsquedas incluyen, además de la supersimetría, otros tipos de nuevas partículas como leptoquarks, o nuevos bosones similares al W y el Z^0 [7]. Téngase en cuenta que la cota que citamos es muy aproximada, ya que

depende, para cada partícula supersimétrica, de sus mecanismos de producción en el acelerador, dentro de la teoría que se pone a prueba. La exploración de masas mayores no se realizará aumentando la energía del acelerador, sino aumentando la intensidad de los haces, ya que sus acoplos disminuyen fuertemente con la masa, en cualquier teoría. Es todavía pronto para poder decir que la supersimetría queda descartada por el experimento.

3) El descubrimiento de que la violación de la simetría materia/antimateria observada en los quarks pesados (b y c) se ajusta bien a la teoría de Kobayashi-Maskawa, sin nuevas fuentes de asimetría. Las medidas realizadas en este sentido por el experimento LHCb han alcanzado gran precisión. Esta teoría se basa en la no universalidad de los acoplos del bosón (o bosones) de Higgs a las distintas generaciones de quarks, cuya jerarquía de masas no es bien comprendida. Sin embargo, los resultados descartan hasta el momento fuentes de nueva física. Al igual que en el caso anterior, el aumento de luminosidad mejorará la precisión en las medidas, y mantendrá el potencial de descubrimiento de acoplos que vayan más allá del Modelo Estándar.

4) El descubrimiento de estados pentaquark de la materia bariónica (similar al protón y al neutrón). El experimento LHCb ha encontrado en 2015 estados cuánticos que se desintegran en protón-charmonium (pJ/ψ) que no pueden ser entendidos como estados de 3 quarks, sino que contienen 5 quarks (uudc \bar{c}). No existían predicciones teóricas específicas para sus masas y números cuánticos. Estos estados exóticos no parecen romper los fundamentos de la teoría de la interacción fuerte de quarks y gluones, pero han supuesto un fuerte revulsivo a la misma, provocando la irrupción de un nuevo tipo de espectroscopía exótica.

5) La observación de algunas anomalías en procesos de desintegración del quark b, que se desvían significativamente del Modelo Estándar en su conjunto, y que aparecen asociadas con la no verificación de la universalidad entre el electrón, el muón y el tau τ [8], que deben tenerlos acoplos de los bosones W y Z^0 en la teoría electrodébil (por la minúscula masa de los neutrinos). Puestas de manifiesto por el experimento LHCb en los últimos años, su estudio ha trascendido a otros experimentos, en particular al experimento Belle en el acelerador KEK de Japón. En el momento presente se considera que los próximos años serán decisivos para dilucidar el origen de estas anomalías, que podría estar tanto en una subestimación de las incertidumbres teóricas y experimentales para fijar la predicción del Modelo Estándar, como en un efecto genuino de nueva física.

6) En lo que respecta a las colisiones de iones pesados, el experimento ALICE, así como su homólogo RHIC en Brookhaven NL, han avanzado significativamente en el conocimiento del plasma de quarks y gluones (QGP), habiéndose establecido que se trata en realidad de un líquido de baja viscosidad, donde los quarks y gluones se encuentran fuertemente acoplados con sus vecinos, en lugar del gas ideal que se pensaba originalmente [9]. El tema tiene gran importancia, porque este líquido es el que llenó el universo en sus primeros microsegundos, dando lugar a los primeros protones y neutrones. Continuarán los estudios en el LHC para caracterizarlo, de manera complementaria a los que se realicen a

temperaturas más bajas, pero con mayor predominio de quarks sobre antiquarks (incluyendo quizá las estrellas de neutrones).

En base a los resultados hasta ahora obtenidos por el LHC, no debe sorprender que el CERN haya aprobado su continuación durante las próximas dos décadas. Esto implicará sucesivas fases de mejora, tanto en la luminosidad del acelerador como en los propios detectores, que se dotarán de las tecnologías más novedosas de instrumentación y de computación.

Auguramos por tanto un futuro brillante para el LHC, como acelerador primordial de nuestro conocimiento de la física de partículas que, no lo olvidemos, guarda muchas de las claves para conocer el universo en que vivimos. Una cuestión que subyace a todas las demás es comprender cómo se transfiere la energía sobre distancias muy cortas, inferiores al tamaño de los núcleos atómicos. Sabemos que la física es profundamente cuántica en este límite, y el LHC nos ha enseñado que el vacío juega en esto un papel esencial. Pero las preguntas se multiplican, e inevitablemente alcanzan los campos de la astrofísica, la física nuclear, la cosmología, y la física de neutrinos. ¿Cuál es el origen de la enorme energía condensada en las masas de algunos quarks? ¿Y de las pequeñas masas de los neutrinos? Es necesario comprender el papel exacto que juega el mecanismo de Higgs en ello. ¿Cuántos bosones de Higgs existen, y cuáles son sus masas? ¿Tienen estructura interna? ¿Cuál es la física que gobierna el valor que esos campos presentan en el vacío, que el Modelo Estándar no nos explica en absoluto? Dado que los quarks pesados no forman parte de la materia ordinaria, que ha dado origen a la vida, ¿cuál es entonces su papel en la naturaleza? Urge saber qué es la materia oscura, y poder aislarla en el laboratorio, sea en un acelerador o en un experimento subterráneo. ¿Existen nuevas partículas pesadas, como los leptarquarks? Y para terminar: ¿qué es la energía oscura observada en astrofísica, y qué relación guarda con el vacío?

El LHC no dará respuesta a todas estas preguntas, pero podría dársela a algunas de ellas. Ante la constatación de que el avance en el conocimiento científico no hace sino generar nuevas preguntas, cada vez más fundamentales, nada mejor que admirar la humilde actitud de Isaac Newton en sus memorias: *“Me veo como un chico jugando a la orilla del mar, que encuentra casualmente una piedra más redonda, o una concha más bonita, mientras el gran océano de la verdad permanece oculto ante mí”*.

Referencias:

- [1] Steven Weinberg, Premio Nobel de Física 1979, <https://www.nobelprize.org/prizes/physics/1979/weinberg/facts/>
- [2] A. Ferrer y E. Ros , ``Física de partículas y de astropartículas``, PUV (2014), <https://puv.uv.es/fisica-de-particulas-y-de-astroparticulas-2a-ed.html>
- [3] B.Adeva, ``La teoría de unificación electrodébil``, Liberlibro (2017), <https://play.google.com/store/books/details?id=pE4sDwAAQBAJ>
- [4] M. Kobayashi y T. Maskawa, Premio Nobel de Física 2008, <https://www.nobelprize.org/prizes/physics/2008/summary/>
- [5] The CMS collaboration, CERN-EP-2018-263, <https://arxiv.org/abs/1809.10733>
- [6] The ATLAS collaboration, ATLAS-CONF-2018-031, <http://cdsweb.cern.ch/record/2629412>
- [7] B2G new physics searches with heavy SM particles (CMS) <https://bit.ly/2RbiBM4>
- [8] P. Koppenburg, Workshop on High-Energy Implications of Flavor Anomalies, Oct. 2018, LHCb-TALK-2018-501, <https://cds.cern.ch/record/2644868/files/Koppenburg-181022CERN.pdf>
- [9] W. Busza, K. Rajagopal y W. dan der Schee, Annual Review of Nuclear and Particle Science Vol. 68:339-376 (Oct.,2018) [10.1146/annurev-nucl-101917-020852](https://doi.org/10.1146/annurev-nucl-101917-020852).

Bernardo Adeva Andany

Catedrático de la Universidad de Santiago de Compostela, y miembro del IGFAE, Instituto Galego de Física de Altas Enerxías, unidad de excelencia María de Maeztu.

Capítulo 20

Medicina personalizada: ¿Presente o Futuro?

La medicina tal y como la hemos conocido está cambiando a un ritmo muy acelerado en las dos últimas décadas. Hasta hace muy poco las enfermedades genéticas se concebían como aquellas entidades en las cuales uno o pocos genes afectos eran capaces de generar un fenotipo que provocaba una enfermedad y la idea era encontrarlo, probar su efecto/defecto y corregirlo. Sin embargo, se sabe que, las cuatro causas de muerte más frecuentes en países desarrollados (enfermedades cardiovasculares, cáncer, accidentes cerebrovasculares y enfermedad pulmonar obstructiva crónica) son todas multigénicas (o sea, todas tienen un componente genético de varios genes afectos) y en las cuales además en medioambiente en el que nos desarrollamos juega un importante papel en el desarrollo y la progresión de la enfermedad. No obstante, la identificación de estos genes y entender el papel que juegan en las enfermedades y cómo y con qué corregirlos no es tarea fácil. Debemos tener en cuenta que nuestro genoma cuenta con 23 cromosomas y contiene la información para alrededor de 30000 genes y la diferencia entre cada uno de nosotros es del 0,1% de toda esta información. Esto quiere decir que buscar entre toda esta información genética aquello que es común entre personas con una determinada enfermedad y que los diferencia de los que no lo están, o (como en el caso del cáncer) encontrar los genes que están diferentes entre las células del tejido normal y el tejido canceroso es una tarea ardua, difícil y que genera una gran cantidad de información que tenemos que ser capaces de procesar para identificar el/los gen/es responsables del fenotipo en cuestión.

En los últimos tiempos nombres como medicina personalizada, medicina de precisión o farmacogenómica se manejan tanto en las publicaciones científicas especializadas como en notas de prensa o noticiarios, pero ¿de qué estamos hablando realmente con esta terminología? Todos estamos familiarizados con la ropa a medida, por ejemplo, pero cuando enfermamos el médico o personal sanitario tiene, no sólo que ser capaz de identificar nuestra dolencia, sino también elegir la correcta medicación para su tratamiento. Pues en este caso, la medicina personalizada se refiere justamente a que en un futuro existan las herramientas clínicas necesarias para que tanto el diagnóstico de cada patología como el tratamiento sea “hecho a medida”.

Según el Instituto Nacional del Cáncer de Estados Unidos de Norteamérica la “*medicina personalizada es el tipo de medicina que usa la información de los genes, las proteínas y el ambiente de una persona para prevenir, diagnosticar y tratar una enfermedad*”. En la medicina personalizada de precisión para el cáncer, se usa información específica del tumor de una

persona con el fin de facilitar el diagnóstico, planificar el tratamiento, determinar si el tratamiento es eficaz o dar un pronóstico (Fig. 1). Ejemplos de medicina personalizada de precisión son el uso de terapias dirigidas para tratar tipos específicos de células cancerosas, como las células de cáncer de mama positivas al gen HER2, y el uso de pruebas de marcadores tumorales para el diagnóstico del cáncer.

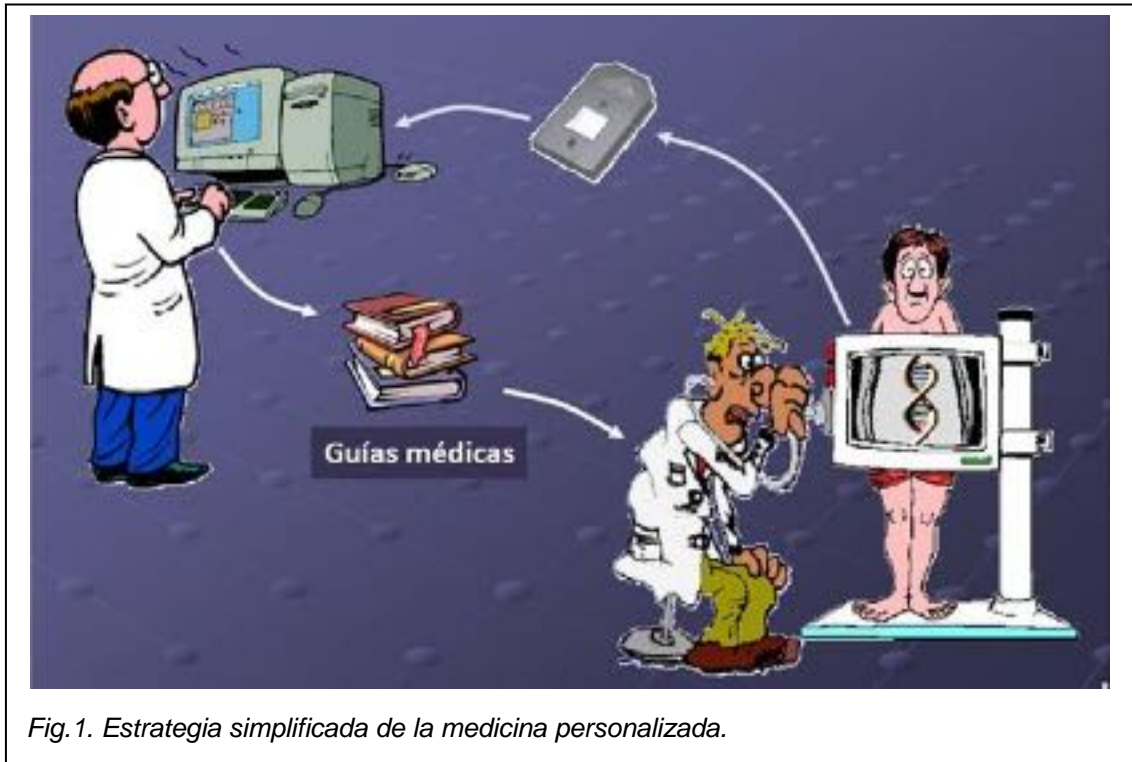


Fig.1. Estrategia simplificada de la medicina personalizada.

La farmacogenómica, sin embargo, es el estudio del modo en que los genes de una persona afectan la manera en que responde a los medicamentos. La farmacogenómica se usa para saber de antemano cuál será el mejor medicamento o la mejor dosis para una persona.

Por lo tanto, la medicina de precisión tiene como objetivo personalizar la atención médica con decisiones y tratamientos adaptados a cada individuo de todas las maneras posibles. La farmacogenómica forma parte de la medicina de precisión, pero no es correcto utilizar ambos términos indistintamente. Aunque el análisis genómico es todavía un desarrollo relativamente nuevo en el desarrollo de nuevos medicamentos, este campo se está expandiendo rápido.

La distinción conceptual entre medicina personalizada y de precisión.

La medicina personalizada se refiere a un acercamiento a los pacientes que considera su composición genética, pero con atención a sus preferencias, creencias, actitudes, conocimientos y contexto social. Mientras que la medicina de precisión describe un modelo para la prestación de atención médica que se basa en gran medida en datos, análisis e información. Este modelo va más allá

de la genómica y tiene vastas implicaciones para la agenda de investigación de cualquier nación y para su implementación y adopción en la atención médica. La medicina de precisión, y el ecosistema que la respalda, deben abarcar el compromiso y la dedicación al paciente, la salud digital, la genómica y otras tecnologías moleculares, el intercambio de datos y la ciencia de la información para que tengan éxito.

Citando palabras de Xabier Pujol Gebelli (Corresponsal científico de El País, y colaborador de revistas como National Geographic y Science&Vie): “No es la primera vez en la historia que se toman decisiones estratégicas basadas en desafíos científicos y tecnológicos de gran calado. Pisar la Luna en menos de una década o la lucha contra el cáncer, fueron dos de las destacadas en el siglo XX; el estudio del cerebro y el desarrollo del grafeno pisan con fuerza en el arranque del XXI. La medicina de precisión, como paso a seguir tras la eclosión del Proyecto Genoma Humano, es la más reciente. Con independencia de su alcance, en todas ellas confluye un factor común: la política y la economía suman con la ciencia como motor”.

Intentaré por tanto dar una visión de conjunto de todos aquellos elementos que, aunque parezcan lejanos de la vida científica, son indispensables para que una revolución tecnológica que es ya una realidad en la biomedicina, se convierta en un beneficio real para toda la población mundial.

Componentes del ecosistema de medicina de precisión.

Un ecosistema de medicina de precisión idealmente vincula a pacientes, proveedores, laboratorios clínicos e investigadores (Fig. 2). Con el advenimiento de los registros médicos electrónicos y los sistemas de desarrollo tecnológico que respaldan tanto la investigación como la prestación de atención médica, los pacientes (y todos aquellos que participan en investigación biomédica voluntariamente), que aceptan proporcionar muestras biológicas y compartir sus datos clínicos y de investigación se encuentran en el epicentro de las contribuciones de esta revolución biomédica.



Fig.2. Ecosistema de la medicina de precisión.

Los investigadores generan nuevos hallazgos a partir de los datos derivados de muestras vinculadas a determinados fenotipos, antecedentes familiares y exposiciones ambientales, todo recabado como parte de la atención clínica. Los médicos utilizan cada vez más conocimientos filtrados a partir de laboratorios clínicos. Este conjunto de datos de una variedad de fuentes establece el escenario para un poderoso ecosistema de medicina de precisión que, cuando se acopla a otros, conduce a una difusión de conocimiento que permite que otros sistemas se beneficien. El ecosistema de medicina de precisión extendida incluye al gobierno como patrocinadores de la investigación de medicina de precisión y reguladores de productos de medicina de precisión, la industria como socios en el desarrollo y comercialización de productos de medicina de precisión, sociedades profesionales como facilitadores de la capacitación de la próxima generación de investigadores, proveedores y analistas políticos que evalúan la idoneidad de las intervenciones de medicina de precisión y el financiamiento que apoya su uso en la atención médica.

El aprendizaje de sistema de salud y la medicina de precisión.

La inclusión de datos genómicos en una infraestructura de sistema de atención de salud que genera conocimiento es una forma de aprovechar todo el potencial de esa información para optimizar la atención al paciente (Fig. 3). En tal sistema, la práctica clínica y la investigación se informan entre sí con el objetivo de mejorar la eficiencia y la eficacia de la prevención, el diagnóstico y el tratamiento.



Fig.3. El aprendizaje de sistema de salud y la medicina de precisión.

Todas las partes interesadas en el ecosistema de medicina de precisión tienen que estar involucradas en la configuración de ese sistema y en la decisión de cómo usar los datos. En particular, los proveedores necesitarán información en el momento de la toma de decisión para poder utilizarla en el

contexto de su flujo de trabajo clínico y los pacientes deberán definir preferencias sobre el uso y el intercambio de su información genómica y de otro tipo con los investigadores. Los investigadores deberán identificar y adoptar las mejores prácticas para la investigación utilizando información genómica vinculada a los registros clínicos informatizados. Los sistemas de salud deberán ofrecer a los proveedores herramientas y sistemas que les permitan tomar decisiones más informadas. La comunidad de tecnología de la información de salud deberá diseñar sistemas seguros e interoperables habilitados para el uso práctico de la genómica y otras nuevas tecnologías, tanto en el ámbito de la salud, como en la comunidad. Los formuladores de políticas sociales deberán abordar el retorno de los resultados, la privacidad, la confidencialidad y la educación, al mismo tiempo que desarrollan regulaciones e incentivos económicos que puedan alinear a todos los interesados hacia los mismos resultados. Los pacientes podrán beneficiarse con resultados de salud optimizados en un sistema de salud de precisión que utilice la información y las tecnologías ómicas como método de aprendizaje y redunde en un mejor tratamiento de las enfermedades.

Ciencia de datos, salud digital y medicina de precisión.

Tres plataformas convergen en la atención de salud: salud digital, ciencia de datos y medicina de precisión. La recopilación a gran escala de conjuntos de datos biológicos, radiológicos y de bases de datos traslacionales se van generando a partir de dispositivos de detección digital e información multiómica (genómica, proteómica, metabolómica, glicómica, etc.) con aplicaciones de apoyo a la investigación y la decisión clínica. El uso completo de estos flujos de datos multidimensionales requiere el desarrollo de métodos estandarizados de adición y análisis de datos y la traducción interdisciplinaria de técnicas computacionales emergentes, como el aprendizaje automático, el procesamiento del lenguaje natural y la inteligencia artificial. La aplicación de estos nuevos métodos analíticos a la atención de salud puede permitirnos definir los patrones dinámicos de salud y enfermedad para crear modelos de atención más eficientes y sostenibles que estén impulsados por datos y tecnología.

El intercambio de datos es una estrategia que merece la pena.

La integración de datos de alta calidad en un sistema de atención médica debe ser una prioridad para garantizar que la mejor información posible esté disponible para la atención e investigación de cada paciente. De hecho, los esfuerzos del gobierno de cada nación y el sector privado deben tener como objetivo establecer bases de datos con información multiómica a gran escala. Sin embargo, actualmente eso aún no está disponible. Los datos tienen que ser estandarizados, comparables y consistentes para facilitar su reutilización en el descubrimiento en múltiples contextos más allá de su uso original. Por ejemplo, las historias clínicas digitales junto con la información de secuenciación génica, ha sido una poderosa herramienta para identificar variantes genéticas asociadas con enfermedades y para comprender la respuesta individual a determinadas terapias. La economía del intercambio de datos podría ser obvia: si un sistema de salud comparte 100 genomas y registros de pacientes con otros 10 que aceptan hacer lo mismo, cada uno gana 900 por el coste de 100.

El poder de la investigación genómica y el cuidado de la salud es proporcional a la cantidad de datos a los que se puede acceder y analizar. Los sistemas de atención médica que soporten ciclos de aprendizaje completos que abarquen tanto el análisis de datos para producir resultados como el uso de esos resultados para desarrollar cambios en las prácticas clínicas serán los sistemas que permitirán un aprendizaje óptimo. Sumados a todos los individuos en tales sistemas, los datos genómicos podrían orientar estrategias para mejorar la salud de la población y contribuir a la mejoría de la atención sanitaria. La información justo a tiempo, las pautas para la acción clínica y más información sobre la utilidad clínica de las pruebas genéticas ayudarían a los médicos a hacer un uso efectivo de la información genómica e integrarla en sus prácticas similares al uso habitual de otras pruebas médicas.

Medicina de precisión - ¿Dónde estamos hoy?

Una serie de aplicaciones de medicina de precisión pueden contribuir al cuidado de la salud en muchos puntos de la vida útil de un individuo. La detección genética se puede usar antes de la concepción para predecir el riesgo de transmitir trastornos genéticos a la descendencia. A las 8-12 semanas de embarazo, una futura madre puede someterse a pruebas genéticas para evaluar las anomalías cromosómicas del feto, e incluso realizar una secuenciación del genoma completo del mismo. Al nacer, la secuenciación se podría utilizar para diagnosticar rápidamente muchas afecciones críticas para las que puede haber resultados procesables que conduzcan a una reducción de la morbilidad y la mortalidad. Más adelante en la vida, estos tipos de enfoques pueden aplicarse para diagnosticar y tratar una variedad de enfermedades, especialmente en el diagnóstico más preciso del cáncer y en la guía de terapias para enfermedades crónicas.

Durante la última década, un número creciente de marcadores genómicos de eficacia, eventos adversos y dosificación de diversos tipos de terapias se han descubierto y recomendado para uso clínico, pero su incorporación a la práctica clínica ha sido variable, incluso cuando se ha respaldado su capacidad de acción con evidencias clínicas. Sin intención de entrar en términos muy específicos, y solo para poner algunos ejemplos me permito incluir casos reales en los que estos marcadores han llevado a un cambio de proceder clínico. En algunos casos, como el genotipo HLA-B * 5701 para el medicamento contra el VIH abacavir, y HLA-B * 1502 para el medicamento anticonvulsivo carbamazepine, los portadores de estos genotipos deben evitar el uso de estos medicamentos para evitar un evento adverso grave. En otros casos, como las variantes en TPMT para mercaptopurina o en CYP2C9 / VKORC1 para warfarina, el ajuste de la dosis del medicamento basado en el genotipo puede ayudar a evitar la toxicidad y mejorar la eficacia.

Una agenda política para la medicina de precisión.

Está claro que la medicina de precisión representa un cambio de paradigma en la atención médica que está madurando y está aquí para quedarse. Ha habido

un rápido aumento en la disponibilidad y el uso de pruebas genómicas y se espera que este crecimiento continúe. Una tendencia particularmente fuerte es el mayor uso de las pruebas multigénicas, incluidos los paneles de genes, la secuenciación completa del exoma y la secuenciación del genoma completo. Para realizar plenamente la integración de la medicina de precisión en la medicina, se debe implementar una agenda de políticas adecuadas a ello. Con el crecimiento en medicina de precisión, surgen oportunidades y desafíos políticos. Un informe de la Academia Nacional de Medicina de los Estados Unidos de Norteamérica señala los siguientes desafíos que utilizaremos para estructurar esta discusión:

1. Generación de evidencia: sigue existiendo la necesidad de evidencia de alta calidad de que la medicina de precisión realmente conlleve a mejorar los resultados para el paciente si se adopta ampliamente. Si bien se han logrado avances en algunas áreas y las iniciativas descritas anteriormente continuarán brindando nuevas pruebas, un desafío importante sigue siendo la evolución aguda y rápida del campo, no solo el crecimiento en la disponibilidad y el uso de tecnologías novedosas en general, sino también el rápido crecimiento de los paneles multigénicos que utilizan tecnologías de secuenciación. Actualmente hay más de 70,000 productos únicos de pruebas genéticas en el mercado y se agregan un promedio de 10 nuevos productos cada día. El mercado de la secuenciación clínica, que abarca el uso de las pruebas de secuenciación para el diagnóstico, la predicción del riesgo de padecer una enfermedad, la selección y el seguimiento de determinadas terapias y el diagnóstico clínico, está creciendo a una tasa anual del 28%. El desafío de la política, por lo tanto, es cómo obtener la evidencia necesaria cuando el campo está creciendo y cambiando tan rápidamente que los "estándar de oro" de los grandes ensayos clínicos aleatorizados puede ser inviable. Se han propuesto varias alternativas que ofrecen enfoques creativos para generar evidencia, tales como nuevos modelos de riesgo compartido y desarrollo de evidencia entre desarrolladores de tecnología y sistemas de atención médica. Desafortunadamente, no habrá un enfoque único para la generación de evidencia, ya que el umbral de la evidencia variará con el riesgo de la prueba y con el impacto financiero en las partes interesadas.

2. Uso compartido de datos y necesidades de infraestructura: la implementación de la medicina de precisión requerirá acceso a datos de pacientes a gran escala, detallados y altamente integrados. Por lo tanto, muchas iniciativas se centran en aumentar la interoperabilidad de los datos del paciente y mejorar los sistemas de manejo de datos que permiten el uso de los datos de la medicina de precisión en el punto de atención primaria. Aunque en los últimos años se han logrado grandes avances hacia el logro de un "sistema de atención médica sin papel" que se basa en los registros médicos electrónicos, se necesita hacer mucho más para integrar los datos en todos los sistemas y para extraer datos que ya existen, pero permanecen en silos. Bajo la rúbrica de "infraestructura" se encuentra el panorama regulatorio. Un tema clave para el futuro será el panorama regulatorio para las pruebas de medicina de precisión. Sigue habiendo incertidumbre sobre cómo los organismos reguladores regularán las pruebas de la medicina de precisión y hasta qué punto aumentarán su supervisión sobre las "pruebas desarrolladas en laboratorio" que no requieren la aprobación de las agencias reguladoras (FDA,

EMA). Estas pruebas requerirán enfoques reguladores flexibles y en evolución paralela a los avances científico-técnicos.

3. Incorporación de datos genómicos y otros datos moleculares en la atención clínica y la investigación: dado que las nuevas innovaciones en la atención de la salud suelen tardar años en ser adoptadas, no es sorprendente que la integración de la medicina personalizada en la atención clínica esté siendo más lenta de lo que algunos observadores han pronosticado. Muchos sectores tienen que unirse para promover su adopción, una educación apropiada, una correcta integración de los sistemas de datos, la dinámica de los procesos del sistema de salud y las políticas de salud. Estos problemas son relevantes para muchos países y regiones que pretenden implementar la medicina personalizada en un futuro próximo. Iniciativas como el Programa de Investigación All of Us en los Estados Unidos y el Proyecto Genómica de los Genomas de Inglaterra (100,000 Genomes) se han difundido ampliamente, pero se ha prestado menos atención a otras iniciativas globales y oportunidades que están ocurriendo fuera de los "epicentros". Por ejemplo, Francia recientemente comprometió \$ 700 millones para financiar centros de secuenciación, y China ha comprometido hasta \$ 10 mil millones para financiar su iniciativa de medicina personalizada. En Europa, el International Consortium for Personalised Medicine (ICPerMed), constituye el proyecto más relevante formado por la Comisión Europea y más de 30 socios europeos y extraeuropeos en representación de Ministerios y agencias financiadoras. Su objetivo fundamental es estimular la investigación y la implementación de la medicina personalizada mediante reuniones, talleres, congresos, encuestas, publicaciones estratégicas e iniciativas conjuntas. Tiene su origen en los talleres preparatorios organizados por la Comisión Europea, junto con varias iniciativas posteriores incluyendo el séptimo Programa Marco y el establecimiento en 2011 de EuroBioForum y el consorcio CASyM. Otros países, como Estonia, Islandia o Reino Unido han desarrollado iniciativas para la creación de biobancos poblacionales que permitan establecer asociaciones entre biomarcadores, historia clínica y estilo de vida. En nuestro país no se ha desarrollado, hasta la fecha, una estrategia general sobre medicina de precisión a nivel estatal. Si bien existen iniciativas como las del Instituto de Salud Carlos III (convocatoria de proyectos en el campo de la medicina de precisión, participación en REDIEX o ELIXIR, etc.); junto con proyectos realizados a nivel local como el Plan Integral sobre Medicina Genómica en Cataluña, el Proyecto Genoma Médico en Andalucía, el proyecto FutureClinic en la Comunidad Valenciana o el proyecto MEDEA de Extremadura. Actualmente en España, específicamente en oncología, se aplican biomarcadores para más de 14 fármacos en más de 14 tumores diferentes, que permiten saber qué pacientes de cáncer de mama, pulmón y colon, entre otros muchos, responderán a un tratamiento específico.

4. Participación y confianza de los participantes: por último, la adopción de la medicina de precisión plantea muchas preguntas sobre el compromiso y la confianza del paciente, que incluyen: ¿Qué constituye realmente? ¿Consentimiento informado? ¿Quién posee información genética y debería tomar decisiones sobre qué resultados se obtienen y cómo se utilizan? ¿Cómo se puede garantizar la privacidad? ¿Qué resultados son los más importantes para los pacientes? Un enfoque del trabajo futuro debe ser el

énfasis creciente en la capacidad de la medicina de precisión para impactar no solo a las personas sino también a las poblaciones, lo que se ha denominado "salud pública de precisión". Se está produciendo el impulso inicial hacia la salud pública de precisión, pero queda mucho por hacer para desarrollar una base probatoria sólida para su uso. Otra área de énfasis debe ser comprender cómo la medicina de precisión podría aumentar o disminuir las disparidades históricas en el acceso a la atención. ¿Las poblaciones históricamente desatendidas tienen acceso a la medicina de precisión y qué políticas podrían garantizar un acceso adecuado? Estas preguntas deberán abordarse dentro del contexto más amplio y cambiante de la reforma de salud mundial.

Un plan de acción para la medicina de precisión.

La realización completa del potencial disruptivo de la medicina de precisión requerirá una agenda científica, clínica y política múltiple. La democratización de los datos sustenta, tanto los avances científicos que permiten no solo la adopción de la medicina de precisión, sino también la medicina *perse*. Se requerirá una cultura con incentivos adecuados para compartir datos. Las partes interesadas del ecosistema de medicina de precisión (participantes, pacientes, proveedores y reguladores) requerirán pruebas de valor en términos de calidad de vida, calidad de atención médica y eficiencia y eficacia optimizadas que justifique el coste. Si tiene éxito, se prestará más atención antes de que se manifieste la enfermedad: se cambiará del tratamiento de la enfermedad a la prevención y detección temprana de la misma. La medicina de precisión requiere liderazgo y perseverancia a nivel mundial para ponerla en el lugar que le corresponde en la salud y la sociedad.

Medicina de precisión: Presente.

La medicina de precisión es una realidad en la práctica clínica asistencial y ha comenzado a cambiar los paradigmas de la medicina e incluso a modificar la forma de clasificación de las enfermedades. Ha dado evidencias fehacientes de que potencia la efectividad y eficiencia ya que permite emplear las estrategias terapéuticas más apropiadas para cada paciente en función del mecanismo molecular subyacente de la enfermedad y las características genéticas individuales. Además, favorece la aplicación del esquema terapéutico más oportuno en los pacientes, ya que toma en consideración la variabilidad genética que determina el metabolismo de los fármacos y su farmacodinámica, junto a factores ambientales que también inciden en su disposición. Así, evita la exposición de los pacientes a fármacos que no son útiles para ellos, reduciendo las posibilidades de efectos adversos relacionados con los medicamentos en pacientes sin posibilidades de respuesta, así como las complicaciones secundarias derivadas de tratar a los pacientes con fármacos ineficaces y la pérdida de oportunidad que ello supone. Desde un punto de vista económico, la medicina de precisión se presenta como una oportunidad de desarrollar un sector industrial de alto valor estratégico, sanitario, científico y económico. Para la correcta puesta en marcha se necesita un proceso de colaboración con el sector farmacéutico y tecnológico y la potenciación de la investigación de

carácter público y su posterior traslación a la práctica clínica. Esta innovación debe suponer un bien común y no un privilegio. Por último, la medicina de precisión se plantea como herramienta para contribuir a la racionalización del gasto sanitario y contribuir a la sostenibilidad del sistema sanitario; generando los datos necesarios para generalizar las iniciativas que demuestren ser coste-efectivas.

Medicina de precisión: Futuro.

Junto a los grandes avances de las investigaciones genómicas que ya se están empleando en tratamientos y estratificación de pacientes oncológicos ya se comienzan a hacer grandes avances en otras patologías que de seguro serán beneficiarias de esta revolución que es ya la medicina personalizada. En unos pocos años, podremos cortar una parte del ADN identificado con ser la causa de una enfermedad y reemplazarla con una secuencia sana. Las nuevas tecnologías como los CrisP-Cas han llegado para quedarse y ampliarán los horizontes de las ciencias ómicas para todas las enfermedades actuales y futuras. Las partes inicialmente vistas como ciencia ficción están, en este mismo momento, convirtiéndose en ciencia basada en hechos y estimulando los próximos 100 años de avances médicos. Pero todavía hay muchas oportunidades para innovar en este campo joven, por lo que es un momento emocionante para participar en la industria de la biotecnología. Los profesionales de hoy brindan soluciones nuevas y vanguardistas a las enfermedades, mejorando así la calidad de vida de la humanidad.

Marta Dueñas Porto

Doctor en Ciencias Biológicas.

Investigadora Titulada Superior en la Unidad de Oncología Molecular (CIEMAT), Madrid.

Capítulo 21

Computación cuántica: un pequeño paso para una partícula, un gran paso para la humanidad.

Los ordenadores son una de las invenciones que más profundamente y con mayor velocidad han transformado nuestro mundo. Aún no se han cumplido cien años del desarrollo de las primeras computadoras y ya están presentes hasta en las tareas más cotidianas. Mientras escribo estas líneas (en un ordenador), escucho música que llega a mis auriculares desde un servidor situado en otro país (después de pasar por varios ordenadores intermedios) a la vez que estoy pendiente del móvil (¡otro ordenador!) para recibir noticias de la revisión médica que tiene hoy mi hija.

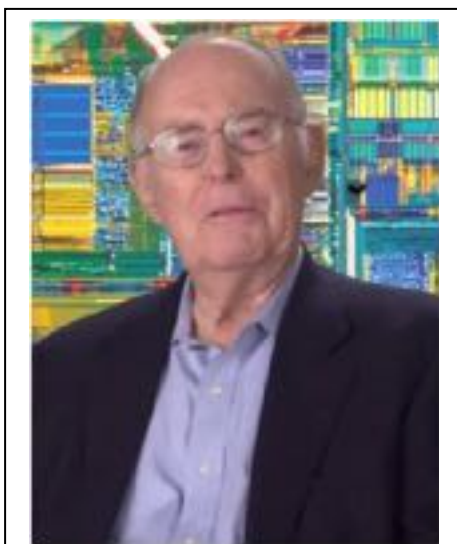


Fig.1. Gordon Moore (Foto de Science History Institute, bajo licencia CC).

Cuando se comenzaron a desarrollar los primeros prototipos, era difícil imaginar el ritmo al que iba a evolucionar la tecnología informática. Sin embargo, ya en 1965 los rápidos avances en la miniaturización de los circuitos integrados llevaron a Gordon Moore [1], uno de los fundadores de Intel, a enunciar su famosa ley⁶⁴: *aproximadamente cada dos años se duplica el número de transistores de los microprocesadores*⁶⁵.

Este crecimiento exponencial ha posibilitado que hoy en día llevemos cómodamente en el bolsillo dispositivos de cómputo que hace apenas unas décadas habrían ocupado habitaciones enteras... si es que hubiera sido posible siquiera construirlos. Pero, de algún modo, la ley de Moore es una *profecía autocumplida*: los fabricantes de circuitos integrados planifican sus estrategias y buscan desarrollar nuevas tecnologías *para que* la ley de Moore se siga cumpliendo.

La pregunta, entonces, es ¿hasta cuándo puede mantenerse este ritmo de crecimiento? Parece evidente que una ley así no puede sostenerse indefinidamente y, de hecho, en los últimos años ya se ha constatado un inicio

⁶⁴ La "ley de Moore" no es una ley en el sentido habitual del término (como puede serlo la ley de la gravitación universal, por ejemplo), sino más bien una observación acerca de la velocidad con la que se producen ciertos desarrollos tecnológicos.

⁶⁵ En realidad, al principio Moore estableció que el periodo necesario para duplicar el número de transistores era de un año, pero posteriormente amplió este tiempo a dos.

de desaceleración [2]. Es más, incluso si se pudiera mantener la tasa de incremento, en algún momento los componentes de los circuitos digitales serían tan pequeños⁶⁶ que se producirían problemas por el efecto túnel cuántico: los electrones podrían “atravesar” espontáneamente la barrera establecida por los semiconductores, impidiendo su correcto funcionamiento.

Con el objetivo de seguir aumentando nuestra capacidad de cálculo, incluso cuando la ley de Moore toque a su fin, se han propuesto numerosas alternativas. Entre ellas, una de las más prometedoras (y de las que más publicidad recibe habitualmente) es la computación cuántica.

Cuando se habla de computación cuántica en los medios de comunicación, se le suelen asociar propiedades cuasimilagrosas (“¡Puede resolver en segundos problemas en los que un ordenador normal tardaría miles de millones de años!”) o que, directamente, parecen magia⁶⁷ (“¡Un *qubit* puede estar en dos estados a la vez!”). Pero ¿cuánto de verdad hay en esas afirmaciones? Y, sobre todo, ¿cómo funciona la computación cuántica y qué la hace tan distinta de la computación tradicional?

En este capítulo, responderemos a esas y otras cuestiones similares, a la vez que intentaremos mostrar brevemente en qué estado se encuentra actualmente el desarrollo de esta tecnología y hacia dónde puede evolucionar en el futuro próximo.

¿Qué *fotones* es la computación cuántica?

Algunas veces, cuando queremos describir de forma rápida lo que es la computación cuántica, decimos simplemente que es el uso de ciertas propiedades de las partículas subatómicas para llevar a cabo cálculos. Sin embargo, esta definición aclara más bien poco y no deja de ser una reescritura poco sutil de las palabras “computación” y “cuántica”. Es más, *toda* la materia obedece las leyes de la física cuántica (y, de hecho, la electrónica moderna está basada en las propiedades de este tipo que presentan los semiconductores) por lo que no parece que esta explicación nos haga avanzar mucho en entender por qué la computación cuántica debería ser distinta de la clásica.

Una descripción más detallada y, seguramente, mucho más certera pasaría por decir que *la computación cuántica explota⁶⁸ propiedades y efectos que solo se manifiestan a niveles microscópicos, como la superposición, el entrelazamiento o la interferencia de las funciones de onda, para llevar a cabo cálculos de forma más eficiente que con los medios de cómputo tradicional.*

⁶⁶ En la actualidad, los procesadores comerciales más avanzados se fabrican con tecnología de 7 nanómetros, un tamaño unas diez mil veces menor que el de un cabello humano.

⁶⁷ Como bien observó el escritor de ciencia ficción Arthur C. Clarke, “cualquier tecnología suficientemente avanzada es indistinguible de la magia”.

⁶⁸ Cuenta Seth Lloyd, uno de los más destacados investigadores en computación cuántica a nivel mundial, que una vez le aconsejaron no utilizar la palabra “explotar” para referirse a este tratamiento de las partículas subatómicas, sino que más bien deberíamos decir que se las “empodera” para realizar cálculos.

Pero claro, ahora hemos transformado una pregunta (¿qué es la computación cuántica?) en otras muchas (¿qué son la superposición, el entrelazamiento y la interferencia y cómo pueden usarse para calcular eficientemente?) de apariencia mucho más técnica y, si se me permite, hasta amenazadora. Sin embargo, quien quiera comprender en profundidad cómo funciona un ordenador cuántico no puede rehuir estas preguntas, por lo que lo explicaremos de la manera más indolora posible. O, cuando menos, lo intentaremos.

Los elementos de la computación cuántica: *qubits*, puertas y medidas.

Describir esquemáticamente el procedimiento que se lleva a cabo para procesar información es una tarea más o menos sencilla. Básicamente, un ordenador recibe unos datos de entrada, que representa internamente de alguna forma, realiza ciertas manipulaciones matemáticas, lógicas, y simbólicas de (la representación de) los datos y, para terminar, ofrece una salida, que es la respuesta que se le da al usuario.

En el caso de la computación tradicional, los datos se representan mediante secuencias de ceros y unos, por lo que se habla de información *binaria* cuya base es el *bit* o *binary digit*. Hay distintas formas de manipular estos bits. Se pueden copiar o mover de un lugar de la memoria a otro, por ejemplo, o se pueden transformar con operaciones que nos permiten sumarlos, multiplicarlos o aplicarles funciones lógicas⁶⁹. Los elementos que realizan estas transformaciones se suelen llamar *puertas lógicas* o *puertas digitales*.

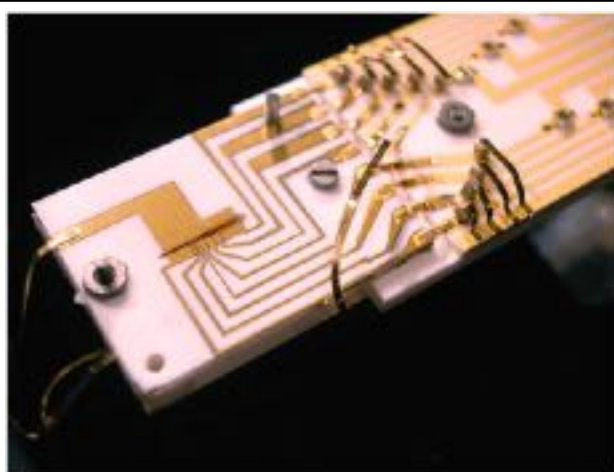


Fig.2. Una trampa de iones, una de las tecnologías utilizadas para implementar físicamente un qubit (Foto de NIST, bajo licencia CC).

Para representar datos más complejos se usan secuencias de varios bits y, si utilizamos varias puertas sobre los bits individuales, podemos obtener las operaciones que deseemos sobre los datos complejos. Los resultados se obtienen al interpretar adecuadamente el patrón de ceros y unos según el tipo de dato que se haya querido representar (un número, un texto, una imagen...).

⁶⁹ En este sentido, se suele identificar el 1 con el valor “verdadero” y el 0 con “falso”. Así, una operación común entre dos bits es el “y lógico” o “and”: el resultado de aplicar un *and* a dos bits es *verdadero* (es decir, 1) si y solo si ambos bits tienen valor *verdadero*.

Aunque parezca imposible, estas sencillas manipulaciones dan lugar a cualquier procedimiento computable que podamos imaginar⁷⁰.

En el caso de los ordenadores cuánticos, hay cambios en todos los elementos del proceso de cálculo: los datos, su manipulación y la forma de obtener la salida. Los datos pasan ahora a identificarse con *qubits* (*quantum bits*), las manipulaciones se llevan a cabo mediante *puertas cuánticas* y para obtener los resultados se han de realizar *mediciones* explícitas. Veamos cada uno de ellos con más detalle.

Los qubits.

A diferencia de los bits, que pueden tomar solo dos valores bien definidos y claramente separados entre sí, los qubits adoptan un rango continuo de valores. Es sumamente común escuchar cosas como que un qubit puede tener el valor 0 y el 1 a la vez, pero (dejando aparte gatos de Schrödinger y demás consideraciones filosóficas del significado de la mecánica cuántica) la afirmación resulta poco precisa, por no decir directamente errónea. Matemáticamente, lo que sucede es que un qubit es una *combinación*⁷¹ de los valores 0 y 1, una especie de valor intermedio entre ellos. Así, podríamos decir que un cierto qubit es un 75% cero y un 25% uno, un 10% cero y un 90% uno, un 57,328% cero y 42,672% uno, o cualquier otra combinación de cero y uno, incluyendo los casos límite en los que es 100% cero o 100% uno. En realidad, la cosa es un poquito más complicada, puesto que los coeficientes de la combinación pueden ser números negativos (e incluso números complejos), pero la idea es esencialmente correcta.

Esta propiedad de los qubits de estar a medio camino entre el 0 y el 1 es lo que se llama *superposición* y es una característica propia del mundo subatómico⁷². Así, cualquier sistema cuántico que pueda encontrarse en dos estados distintos (por ejemplo, un fotón con dos direcciones posibles de polarización o una partícula con *spin*, que puede estar “girando” en dos direcciones diferentes) se presenta, potencialmente, como una posibilidad para implementar físicamente un qubit.

Una representación gráfica habitual de un qubit es la llamada *esfera de Bloch* (ver Fig. 3). En ella, los números que definen a un qubit como combinación del 0 y el 1 se traducen en dos coordenadas que lo sitúan en un punto de la superficie de una esfera, con el 0 situado en el polo norte y el 1, en el polo sur⁷³. Nótese que, mientras que en computación clásica se trabaja solo

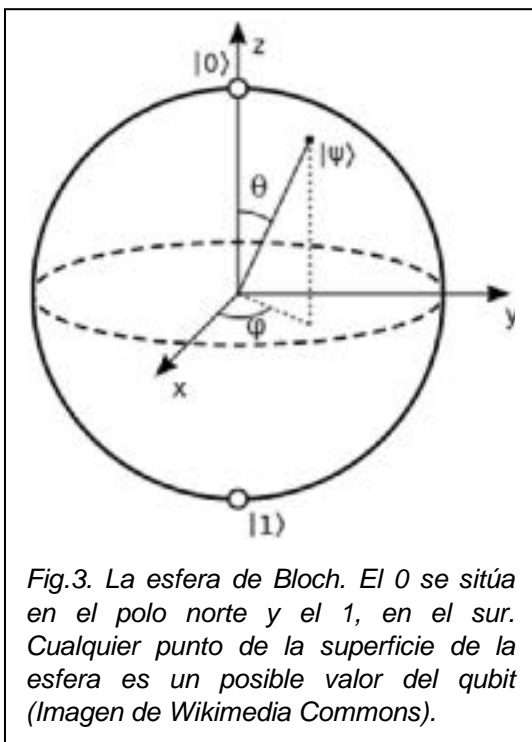
⁷⁰ Este hecho suele llamarse, técnicamente, la tesis de Church-Turing.

⁷¹ Formalmente, un qubit es un elemento de un espacio vectorial de Hilbert de dimensión dos, por lo que es una combinación lineal de dos elementos de una base ortonormal del espacio. Los elementos de la base se suelen denotar como $|0\rangle$ y $|1\rangle$.

⁷² La ecuación de Schrödinger, a la que obedecen los sistemas cuánticos, es lineal, lo que quiere decir que la superposición (es decir, la suma ponderada) de varias soluciones de la ecuación es también una solución de la ecuación.

⁷³ En realidad, hay otro parámetro extra, una fase global que puede afectar a todo el qubit y no queda reflejada en esta representación, pero este valor no influye en las mediciones y, por tanto, podemos obviarlo en este punto.

con los dos valores 0 y 1, en computación cuántica se usan un número infinito de posibilidades puesto que cualquier punto de la esfera representa un posible qubit.



El potencial del uso de qubits en lugar de simples bits clásicos se comienza a entrever cuando manejamos no uno, sino varios qubits. En ese caso, un estado del ordenador cuántico es, en general, una superposición de *todas* las posibles configuraciones de ceros y unos de los qubits individuales⁷⁴. Si, por ejemplo, manejamos tres qubits, tendremos que el estado es una suma de las combinaciones 000, 001, 010... hasta 111 (ocho posibilidades distintas). Cuando tenemos n qubits, el número de configuraciones es 2^n y un solo estado puede ser una superposición de todas ellas a la vez. De este modo, cuando hacemos una operación sobre un estado de ese tipo, de algún modo estamos operando sobre los 2^n valores. La mayor parte de los algoritmos

cuánticos explotan esta propiedad para acelerar los cálculos con respecto a los algoritmos clásicos (pero, como veremos, no es una cuestión trivial y no se puede conseguir en todas las situaciones).

Las puertas cuánticas.

Ahora ya sabemos que en un ordenador cuántico los datos se representan internamente mediante qubits, pero ¿cómo se manejan? Puesto que cada conjunto de qubits en un estado concreto es, en realidad, el estado físico de un cierto sistema cuántico, las transformaciones a las que podemos someterlo son aquellas permitidas por la mecánica cuántica. En concreto, un sistema cuántico queda caracterizado por una función, llamada función de onda, que nos dice en cada instante qué valores tienen los coeficientes de la superposición. Esta función de onda evoluciona según la famosa ecuación de Schrödinger⁷⁵ y, por tanto, las transformaciones permitidas son lo que se llaman *transformaciones unitarias*⁷⁶.

⁷⁴ Matemáticamente, tenemos un producto tensorial de los espacios (bidimensionales) donde viven cada uno de los qubits individuales.

⁷⁵ La ecuación de Schrödinger es una ecuación diferencial que especifica cómo evoluciona un sistema cuántico.

⁷⁶ En el caso de un número finito de dimensiones, una transformación unitaria viene dada por una matriz cuadrada de números complejos cuya inversa es, precisamente, la matriz que se obtiene tomando la transpuesta y cambiando cada número por su conjugado complejo.

En el caso de trabajar con un solo qubit, las transformaciones unitarias se pueden representar como rotaciones de la esfera de Bloch y, en general, incluso cuando trabajamos con más de un qubit, las transformaciones permitidas son rotaciones de todo el espacio en el sentido de que preservan los ángulos entre las distintas superposiciones de qubits. Esto tiene varias consecuencias importantes, puesto que marca una clara diferencia con respecto a la computación clásica.

Por un lado, todas las operaciones cuánticas son reversibles (puesto que son “rotaciones”, siempre se puede volver “hacia atrás” rotando en sentido contrario). Esto implica que hay ciertos procesos que no se pueden realizar. Por ejemplo, en computación clásica podemos borrar el contenido de una posición de memoria y ponerlo a cero. Sin embargo, esto es imposible en computación cuántica. Imaginemos, por ejemplo, que tenemos almacenado el valor 01 mediante un par de qubits y el valor 10 en otros dos. Si borramos cada uno de ellos, obtendremos 00 en ambos casos. Como las transformaciones cuánticas son reversibles, deberíamos poder volver desde 00 a 01 y a 10, respectivamente, pero al borrar hemos perdido la información que nos permitía deshacer la operación. ¿Debemos volver desde 00 a 01 o a 10? Es imposible saberlo y, por tanto, la operación de borrado no es físicamente realizable⁷⁷.

Otra operación habitual en la computación clásica y que, curiosamente, no puede realizarse en computación cuántica es la copia de información. Se puede demostrar con bastante facilidad (es el famoso *Teorema de no clonación* [3]) que no es posible hacer, en general, una copia independiente del estado de un grupo de qubits en otro grupo de qubits diferente⁷⁸.

Pese a estas limitaciones, que pueden parecer muy severas para alguien acostumbrado a las técnicas de programación clásicas (donde copiar y borrar son operaciones ubicuas), es relativamente fácil demostrar [4] que todo lo que se puede realizar con bits y puertas lógicas clásicas se puede conseguir también con qubits y una adecuada elección de las transformaciones unitarias (por ejemplo, con las llamadas puertas de Toffoli y de Fredkin). Eso sí, el coste añadido es que se necesitarán qubits auxiliares para guardar todos los resultados intermedios, ya que es imposible borrar información⁷⁹.

Otra diferencia aún más notable entre las transformaciones que se pueden realizar en computación cuántica y las de la computación clásica es que mientras que en este último caso el número de puertas diferentes para un número fijo de bits es siempre finito, en la computación cuántica el número de posibles operaciones (incluso con un solo qubit) es infinito (puesto que hay infinitas formas de rotar la esfera de Bloch). Obviamente, cuando queremos

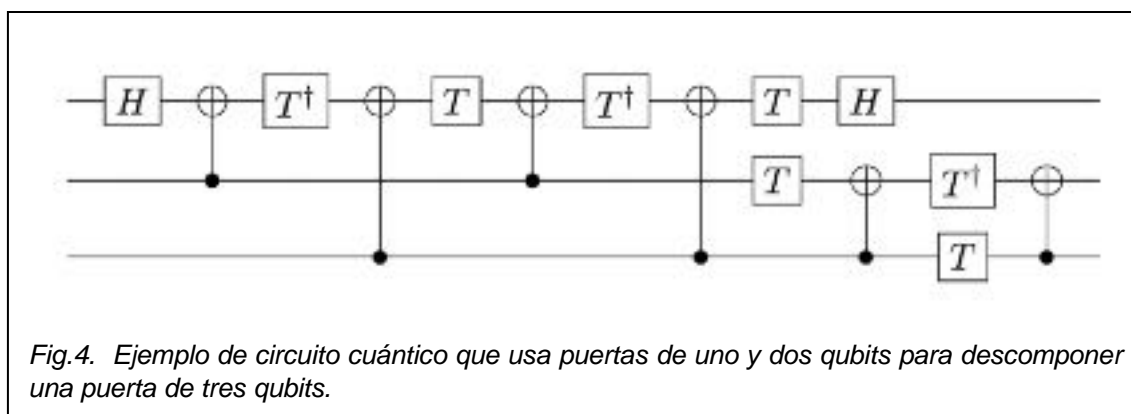
⁷⁷ ¿Cómo es posible, entonces, que un ordenador clásico borre información si, en realidad, también sigue las leyes cuánticas? La respuesta es que, en realidad, “borra” esa información gastando energía y transfiriéndola al entorno, aumentando de esta manera su entropía. La información sigue estando ahí, pero es irre recuperable en la práctica.

⁷⁸ Lo que sí se puede hacer es una copia *dependiente*, en el sentido de que ambos grupos de qubits quedarán entrelazados y, en lo subsiguiente, cualquier operación que se realice sobre uno de los dos grupos afectará también al otro.

⁷⁹ Hay un pequeño truco, sin embargo, que permite “descomputar” algunos valores y volver a reutilizar los qubits auxiliares.

escribir un algoritmo para ejecutar en un ordenador cuántico, no disponemos de un número infinito de operaciones distintas así que ¿qué podemos hacer?

La solución pasa por utilizar ciertos resultados matemáticos⁸⁰ que permiten asegurar que usando sólo un número finito de operaciones de uno y dos qubits es posible *aproximar* cualquier transformación unitaria que se desee con un error tan pequeño como se requiera⁸¹. Este es el camino que se sigue en la mayor parte de los ordenadores cuánticos que se están construyendo en la actualidad, en los que se selecciona un número relativamente pequeño de operaciones con un número reducido de qubits. Estas operaciones reciben el nombre de *puertas cuánticas*, por analogía con las puertas lógicas, y se usan para construir (o aproximar, según el caso) todas las demás que se necesiten. Estos grupos de puertas cuánticas reciben el nombre de *circuitos cuánticos* (ver Fig. 4).



Entre las puertas más usadas en casi todas las arquitecturas se encuentran las rotaciones de un solo qubit y la puerta CNOT, que actúa sobre dos qubits⁸². La puerta CNOT permite realizar *entrelazamientos* entre dos qubits, de manera que el valor de uno queda ligado al de otro de un modo que es exclusivo del mundo cuántico. Por medio de la puerta CNOT, podemos construir superposiciones de dos qubits que son un 50% 00 y un 50% 11. Si medimos uno de los qubits de este tipo de estados, obtendremos con el 50% de probabilidad un 0 y con el 50% de probabilidad un 1. Supongamos que obtenemos un 0. Entonces, el estado “colapsa” y queda transformado en 00. Si a continuación medimos el otro qubit, obtendremos 0 con el 100% de probabilidad. Con el caso 11 sucede exactamente lo mismo. Así que midiendo estos estados entrelazados obtenemos valores completamente aleatorios en cada qubit por separado pero, a la vez, el valor de ambos qubits es el mismo. Este hecho tiene importantes consecuencias y aplicaciones, por ejemplo, en criptografía cuántica (de la que hablaremos brevemente hacia el final de este capítulo).

⁸⁰ Véase, por ejemplo, [5].

⁸¹ Pero, en general, ese error existirá y deberá tenerse en cuenta al analizar el funcionamiento de nuestro algoritmo.

⁸² La puerta CNOT es una negación controlada: cambia el valor del segundo qubit de 0 a 1 y de 1 a 0, pero sólo si el primer qubit tiene valor 1; si el primer qubit es 0, el segundo qubit permanece inalterado.

Otra puerta muy importante es la puerta de Hadamard o puerta H, que permite poner un qubit en una superposición que es mitad 0 y mitad 1. De esta forma, aplicando puertas H a cada uno de los qubits de un grupo, se puede obtener una superposición con igual contribución de cada una de las posibles configuraciones de ceros y unos, desde 00...0 hasta 11...1. Muchos algoritmos cuánticos comienzan construyendo una superposición de este tipo, puesto que de esta forma las operaciones posteriores se aplican sobre todas esas configuraciones a la vez, aprovechando lo que se suele llamar el *paralelismo cuántico*.

Las medidas.

Con lo explicado hasta ahora, podría dar la impresión de que aprovechar la computación cuántica para realizar cálculos de forma mucho más rápida que con un ordenador clásico es pan comido. Si, por ejemplo, quisiéramos evaluar una función en un montón de valores a la vez, parecería que sería suficiente con construir un circuito cuántico para esa función y aplicarlo a una superposición de qubits tal como se explica en el último párrafo de la sección anterior. De esta forma, se calcula la función en todos los valores desde el 00...0 hasta el 11...1 de una sola vez, en contraste con las 2^n aplicaciones de la función que serían necesarias en un ordenador clásico.

Esta idea es esencialmente correcta y, de hecho, se encuentra en el núcleo de gran parte de los algoritmos cuánticos conocidos, pero tiene un pequeño problema: cómo recuperar la información.

En un ordenador clásico, consultar el resultado de una operación es tan sencillo como observar los valores que se encuentran almacenados en una serie de bits de la memoria. Pero recordemos que en un ordenador cuántico los datos se almacenan en qubits que normalmente se encuentran en una superposición de estados. Esa superposición cuántica no se puede observar directamente, sino que es necesario realizar una medida que arroja como resultado una cadena de bits y que, además, destruye de forma irreversible la superposición.

Por ejemplo, si tenemos un qubit en un estado que es mitad 0 y mitad 1, cuando lo midamos obtendremos como resultado 0 el 50% de las veces y 1 el otro 50%. Además, a partir de ese momento el qubit pasará a estar en estado 0 o 1, según el resultado obtenido, ya sin superposición alguna.

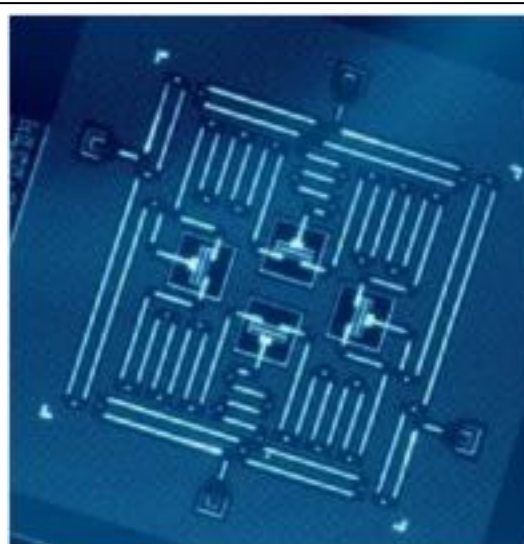


Fig.5. Chip cuántico de IBM, con 4 qubits (Imagen de Jay M. Gambetta, Jerry M. Chow y Matthias Steffen, bajo licencia CC).

Por este motivo, la estrategia esbozada al inicio de esta sección no se puede usar directamente: si aplicamos una función sobre una superposición, el estado cuántico será una superposición de todos los valores de la función a la vez, pero al intentar “leerlos”, obtendremos solo uno de ellos... y además será uno al azar, no el que nosotros queramos.

El arte de programar un ordenador cuántico pasa, por tanto, por ser capaz de encontrar operaciones que permitan ir modificando nuestra superposición de manera que aumente la probabilidad de obtener el resultado que necesitamos mientras disminuye la de los valores no deseados. Por ejemplo, si tenemos una superposición en la que el 00 es el valor “correcto” e inicialmente su proporción es del 25%, nuestro circuito deberá localizarlo haciendo que aumente su probabilidad hasta un valor mucho más alto (un 90%, por ejemplo), para que al medir lo obtengamos en la mayor parte de los casos. Si esa probabilidad es suficientemente alta, quizá no hallemos el valor a la primera, pero repitiendo todo el proceso un número pequeño de veces lo obtendremos casi con total seguridad.

Los siguientes apartados describirán a vista de pájaro cómo se pueden implementar este tipo de ideas para resolver ciertos problemas de forma más rápida que en un ordenador clásico.

Un problema que no interesa a nadie... o puede que sí.

Hasta ahora hemos descrito los distintos elementos que se usan para desarrollar circuitos cuánticos (qubits, puertas cuánticas y medidas) y, ¡por fin!, estamos en condiciones de describir algunos de los algoritmos más famosos de la computación cuántica. Comenzaremos con uno de los más sencillos de comprender, un problema que no tiene demasiado interés práctico pero que fue el primero para el que se desarrolló un método cuántico eficiente: el algoritmo de Deutsch-Jozsa [6].

Para comprender el problema, imaginemos que tenemos un armario con cuatro cajones, numerados del 0 al 3. En cada uno de ellos hay una bola, que puede ser de color blanco o de color negro. No sabemos de qué color es la bola de cada cajón antes de abrirlo, pero nos han prometido que o bien todas las bolas son del mismo color o bien la mitad son de un color y la otra mitad de otro. Es decir, o hay cuatro bolas blancas, o hay cuatro bolas negras o hay dos blancas y dos negras. El problema consiste en determinar con total certeza en cuál de las dos situaciones estamos (todas las bolas iguales o la mitad de cada color) *abriendo el menor número posible de cajones*.

Es obvio que abrir un solo cajón no sirve para resolver el problema, independientemente de cuál sea el color de la bola que nos encontremos. Si abrimos un segundo cajón, hay dos posibilidades: o bien la bola es del mismo color que la del primer cajón que abrimos o bien es de color opuesto. En este segundo caso, ya no necesitamos abrir más cajones, porque sabremos inmediatamente que no todas las bolas son del mismo color y, por tanto, la mitad serán blancas y la mitad negras. Si hemos obtenido bolas del mismo color en ambos cajones, necesitaremos abrir un tercero y, ahí sí, salga lo que

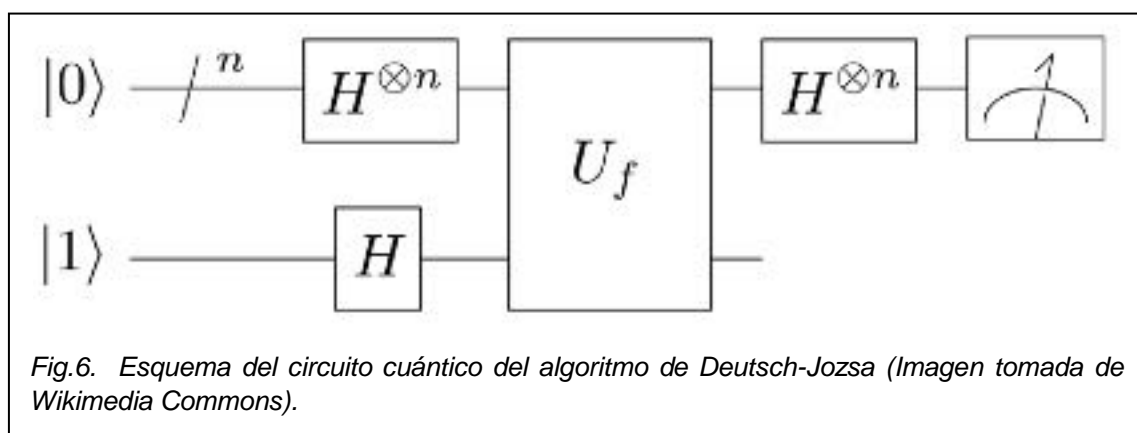
salga, habremos resuelto el problema: si la tercera bola también es del mismo color, entonces la que hay en el cajón sin abrir también será igual (recordemos que nos prometieron que serían todas iguales o la mitad de cada color, y ya hemos visto tres del mismo color); si, por el contrario, ahora obtenemos una bola distinta, entonces no todas son del mismo color y la única posibilidad que nos queda es que sean blancas y negras a partes iguales.

Pensemos ahora en el mismo problema, pero con cajones numerados del 0 al 2^n-1 . De nuevo, nos aseguran que o bien todas las bolas son del mismo color o bien hay una mitad de cada color. ¿Cuántos cajones deberemos abrir para averiguar en cuál de las dos situaciones nos encontramos? Si tenemos suerte y encontramos inmediatamente dos bolas distintas, con abrir dos cajones será suficiente. Pero si todas las bolas son del mismo color, nunca podremos estar completamente seguros hasta que no abramos $2^{n-1}+1$ cajones (la mitad más uno).

Pues bien, podemos plantear este problema como una cuestión puramente matemática y resolverlo mediante computación cuántica... ¡abriendo solamente un cajón!

En lugar de cajones y bolas blancas y negras, supondremos que tenemos una función f que a cada valor del 0 al 2^n-1 le asigna un 0 o un 1. Entonces podemos identificar la afirmación “hay una bola negra” en el cajón x con “el valor de $f(x)$ es 1”, y “hay una bola blanca” en el cajón x con “el valor de $f(x)$ es 0”. Así, “abrir el cajón x ” es, ahora, “consultar el valor $f(x)$ ”. Por la suposición acerca de las bolas que hay en los cajones, la función será o bien *constante* (vale siempre 0 o siempre 1) o bien *balanceada* (vale 0 en la mitad de los “cajones” x y 1 en la otra mitad).

La función f se puede implementar mediante puertas cuánticas para usarla en un circuito que nos permita determinar en cuál de las situaciones estamos (ver Fig. 6). Basta con crear una superposición de los valores del 0 al 2^n-1 , evaluar la función f sobre todo ellos a la vez y, con un pequeño truco, conseguir que cuando $f(x)$ sea 1, el coeficiente correspondiente a x cambie de signo. Finalmente, deshacemos la superposición y medimos.



Si todos los valores de f son iguales, no pasará nada y obtendremos 00...0 como resultado. Sin embargo, si la mitad de los valores de f son 0 y la otra mitad son 1, el signo negativo de unos se compensará con el positivo de

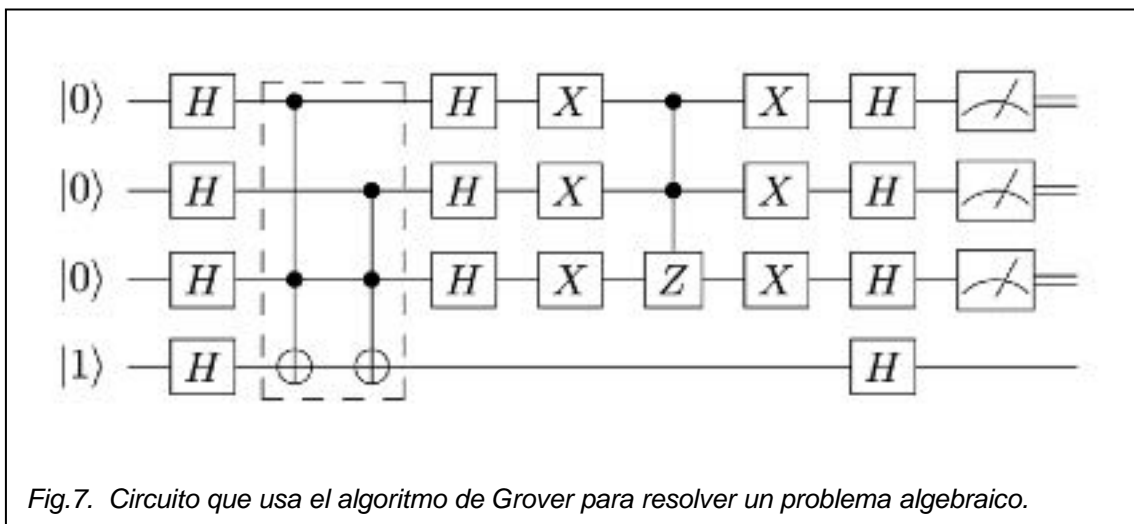
otros (decimos que se produce una *interferencia destructiva*) y obtendremos un valor distinto de 00...0. Así, podemos fácilmente saber en qué situación nos encontramos... consultando f una sola vez.

Aunque el problema es poco útil en la práctica, hemos demostrado algo muy importante: con un ordenador cuántico podemos determinar una propiedad de la función f mediante *una sola evaluación*, mientras que con un ordenador clásico necesitaríamos, en el peor caso, aplicar f $2^{n-1}+1$ veces.

Buscando una aguja en un pajar.

Un ejemplo más útil de lo que un ordenador cuántico puede llegar a lograr lo encontramos en el llamado algoritmo de Grover [7]. El problema que debemos resolver en este caso es mucho más natural: tenemos una lista de elementos y queremos encontrar uno que cumpla una condición concreta. Por ejemplo, imaginemos que tenemos un listín telefónico y queremos averiguar el nombre de la persona a la que corresponde un cierto número. Como la lista no está ordenada por números, con un ordenador clásico nuestra única opción es ir mirando uno por uno todos los teléfonos hasta encontrar el que nos interesa.

El algoritmo de Grover nos permite reducir el número de veces que tenemos que consultar la lista de elementos. En lugar de tener que mirar los N registros, podemos mirar solo \sqrt{N} con la garantía de tener una probabilidad muy alta de encontrar el elemento que buscamos (recordemos que, en general, los algoritmos cuánticos no son deterministas y podría ser necesario repetirlos varias veces para tener éxito). Por ejemplo, si nuestra lista tuviera 10.000 elementos, con el algoritmo de Grover solo sería necesario consultarla unas 100 veces para dar con el dato que necesitamos.



Este método comparte algunas ideas con el algoritmo de Deutsch-Jozsa. Se forma una superposición de todos los elementos y se evalúa sobre todos ellos una función que cambia el signo de aquel que nos interesa encontrar. A continuación, se hace una rotación que va ampliando poco a poco la probabilidad del valor buscado. Este proceso se repite aproximadamente \sqrt{N}

veces y, voilá, cuando midamos tendremos muchas posibilidades de hallar el elemento buscado.

El tiempo ganado no es tan espectacular como en el caso del problema de los cajones y las bolas, pero a cambio es un algoritmo con muchísimas más aplicaciones prácticas y que, por ejemplo, puede utilizarse también para encontrar mínimos o máximos [8].

Rompiendo códigos con un ordenador cuántico.

Sin duda el algoritmo que más fama le ha dado a la computación cuántica es el de Shor [9], que permite descomponer un número entero en sus factores primos. Este proceso, que a primera vista parece carecer de utilidad práctica alguna, es de vital importancia por su relación con algunos algoritmos criptográficos muy utilizados, como el famoso RSA [10].

La fortaleza de estos métodos de cifrado de información reside en que, a día de hoy, no conocemos ningún algoritmo que se pueda ejecutar eficientemente en un ordenador clásico y que sea capaz de encontrar los factores primos de un número entero dado. En el caso del RSA, por ejemplo, cada usuario tiene una clave pública (que se puede usar para enviarle mensajes que nadie más pueda leer y para certificar los mensajes que ha firmado) que es el producto de dos primos muy grandes⁸³. Cada parte debe mantener sus números primos en secreto, porque si se conocieran, se podrían utilizar para leer sus mensajes y para suplantar su identidad.

La importancia del algoritmo de Shor radica en que se puede utilizar para obtener esos números primos de manera mucho más rápida que con un algoritmo tradicional. El mejor algoritmo clásico conocido⁸⁴ emplea un tiempo que crece (sub-)exponencialmente con el número de dígitos del entero a factorizar. En la práctica, esto supone que aumentar en un solo bit la longitud de la clave multiplica por un número mayor que uno el tiempo que se tarda en romperla. Así, aumentar en, digamos, 100 bits el tamaño de la clave, cosa relativamente sencilla, multiplicaría por un número inmenso el tiempo necesario para factorizarla.

Sin embargo, el algoritmo de Shor emplea un tiempo que sólo crece polinomialmente⁸⁵, con lo que un aumento de la longitud de la clave se traduce en un aumento moderado (al menos en comparación con el caso clásico, ver Fig. 8) del tiempo empleado en romperla. Explicar los detalles del proceso requeriría bastante más espacio del que tenemos disponible, así que baste con decir que la clave de la eficiencia del algoritmo de Shor se basa en la posibilidad de ejecutar mucho más rápidamente que en un ordenador tradicional el cómputo de la transformada discreta de Fourier, que permite hallar el periodo de ciertas funciones aritméticas.

⁸³ Cada uno de ellos con cientos de dígitos.

⁸⁴ Basado en la criba general del cuerpo de números (*general number field sieve*).

⁸⁵ Con más precisión, tiene orden de crecimiento $O(n^2 \log(n) \log(\log(n)))$, siendo n el número de dígitos de la clave.

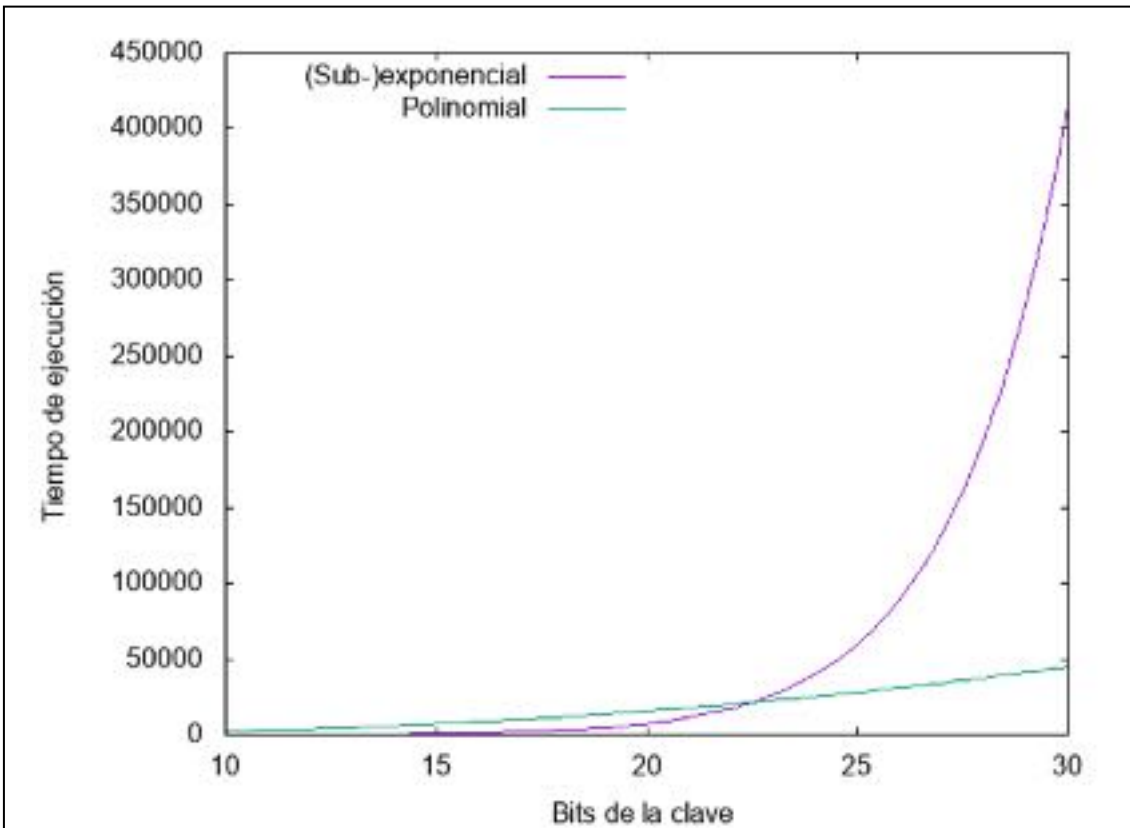


Fig.8. Comparación entre el tiempo de ejecución de un algoritmo subexponencial y de uno polinomial en función de la longitud de la clave a factorizar. A partir de un cierto valor, el algoritmo polinomial es mucho más rápido.

¿Qué más pueden hacer los ordenadores cuánticos?

Además de la posibilidad de ejecutar los algoritmos ya mencionados (Deutsch-Jozsa, Grover y Shor), los ordenadores cuánticos abren otras posibilidades para realizar más eficientemente tareas que son costosas (o prácticamente imposibles) con recursos tradicionales.

Por ejemplo, ya en 1982, Richard Feynman observó [11] que simular la evolución de un sistema cuántico es un proceso intratable en un ordenador clásico. Como ya hemos mencionado, el número de parámetros necesarios para especificar un estado cuántico de n qubits es 2^n . Si n es del orden de varios cientos o miles, sólo la cantidad de bits de memoria necesarios para guardar el estado superaría el número de partículas de todo el universo. Sin embargo, los ordenadores cuánticos *son* sistemas cuánticos, por lo que es natural utilizarlos para reproducir el comportamiento de otros sistemas del mismo tipo. Esto abre la posibilidad, entre otras muchas cosas, de estudiar reacciones químicas complejas, con el abanico de aplicaciones que eso supone, desde el diseño de nuevos materiales al desarrollo de fármacos más efectivos.

Además, en ciertas situaciones, es posible resolver sistemas de ecuaciones lineales eficientemente en ordenadores cuánticos [12]. Por medio

del algoritmo HHL⁸⁶, podemos resolver este tipo de problemas de forma exponencialmente más rápida que con los métodos clásicos. La limitación del proceso radica en que los datos de entrada deben estar codificados como estados cuánticos y la salida también se da como un estado cuántico. Aunque esto puede impedir recuperar la información de forma rápida, como ya hemos visto, se han encontrado aplicaciones del método HHL para acelerar algoritmos de inteligencia artificial y aprendizaje automático (*machine learning*). De hecho, el campo del *quantum machine learning* es uno de los más prometedores dentro de la investigación actual en computación cuántica.

Otro ámbito que puede beneficiarse ampliamente de los desarrollos en las tecnologías cuánticas de la información es el de la criptografía. No se trata de una aplicación pura de la computación cuántica, pero está relacionada de manera muy íntima con ella. La idea es usar propiedades como la superposición y el entrelazamiento de qubits para desarrollar esquemas (uno de los más conocidos es el BB84 [13]) que permiten compartir claves cien por cien seguras y, a la vez, cerciorarse de que nadie ha estado espiando el intercambio de información. La clave es que al medir dos bits entrelazados se obtienen valores aleatorios pero idénticos, como ya hemos visto. Si Alice y Bob quieren comunicarse de forma secreta, pueden compartir pares de qubits entrelazados y utilizarlos para intercambiar una clave de cifrado que se puede demostrar que es irrompible. Así que, ironías de la vida, aunque las tecnologías cuánticas pueden servir para romper la criptografía tradicional, a cambio podrían ofrecernos un medio incluso más seguro de intercambio de claves y, por tanto, de comunicación.

Y todo esto, ¿para cuándo?

Está muy bien hablar de algoritmos y de sus potenciales aplicaciones, pero de nada sirve un método que funciona de forma perfecta sobre el papel si no tenemos un ordenador donde ejecutarlo. La pregunta del millón es, por tanto, cuándo podremos disponer de ordenadores cuánticos que se puedan utilizar en la práctica. Y la respuesta es que... no resulta fácil decirlo.

Aunque en los últimos años se está impulsando de forma significativa la investigación en tecnologías cuánticas y son muchas las grandes compañías, como IBM, Google o Microsoft, que están invirtiendo decididamente en su desarrollo, lo cierto es que los retos a superar no son triviales. Como hemos visto, la mayor parte de los algoritmos dependen de la posibilidad de poner nuestros qubits en superposición, de realizar sobre ellos operaciones que los modifiquen y los entrelacen, y de mantener estos estados durante periodos de tiempo lo suficientemente largos como para completar los cálculos.

El problema es que estos sistemas son sumamente delicados y cualquier interacción con el entorno les hace perder las propiedades que necesitamos para nuestros cálculos⁸⁷. Se están estudiando y probando distintas propuestas para la implementación de qubits y puertas cuánticas (por

⁸⁶ Sus descubridores son Aram Harrow, Avinatan Hassidim y Seth Lloyd.

⁸⁷ Técnicamente, decimos que se pierde la coherencia de los estados cuánticos.

ejemplo, mediante trampas de iones o con circuitos superconductores) pero en casi todas ellas es necesario mantener temperaturas muy cercanas al cero absoluto⁸⁸ así como un grado extremo de aislamiento ante cualquier perturbación externa.

En los últimos años, algunos de estos ordenadores han pasado de ser prototipos de laboratorio a convertirse, como en el caso de los fabricados por IBM, en dispositivos accesibles de forma gratuita y libre a través de la red [14], posibilitando que cualquier persona interesada pueda realizar experimentos con ellos. Por el momento, sin embargo, el número de qubits de estas computadoras es del orden de unas pocas decenas⁸⁹ y las puertas cuánticas aún presentan más error del que sería necesario para realizar cálculos largos y estables. Se habla, por tanto, de tecnología cuántica de escala intermedia y con ruido⁹⁰ que, aun así, puede ser útil en ciertas tareas que son difíciles para los ordenadores tradicionales [15].

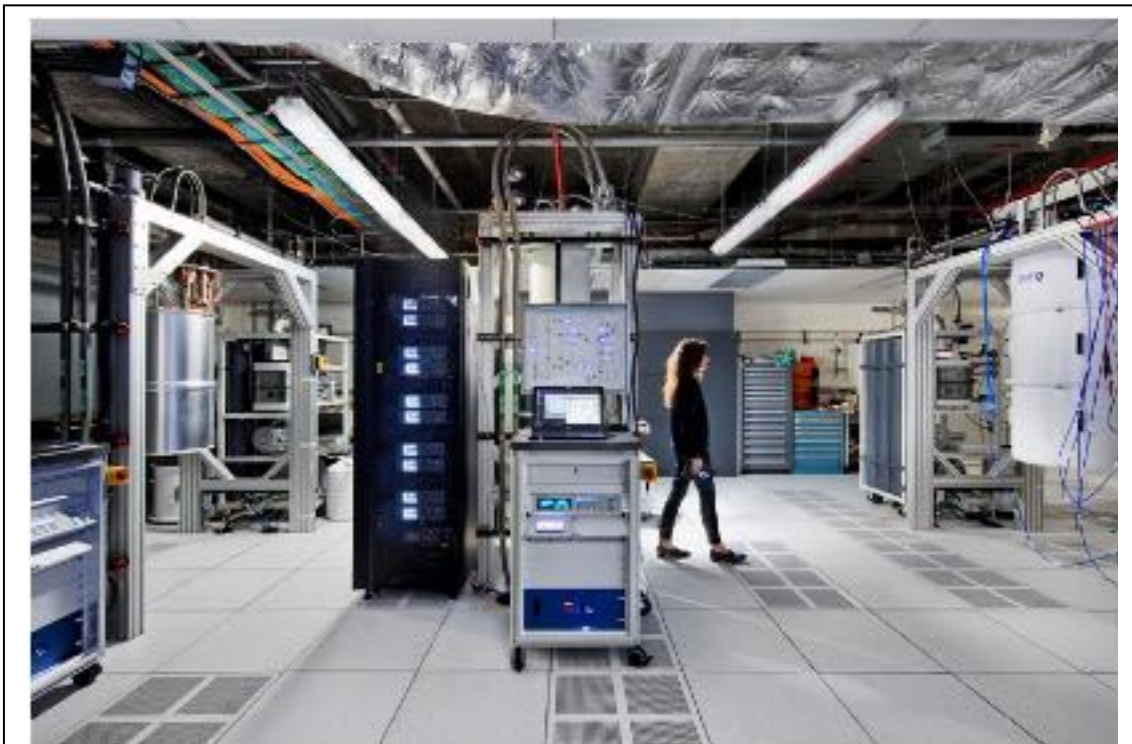


Fig.9. Laboratorio con varios de los ordenadores cuánticos de IBM (Foto de Connie Zhou para IBM).

Cuando todas estas tecnologías maduren, algo que según algunos expertos podría ocurrir tan pronto como en unos diez años [16], lo más probable es que los ordenadores cuánticos no reemplacen por completo a los computadores clásicos, sino que convivan con ellos, especializándose en las

⁸⁸ Tan bajas, en algunos casos, como 20 milikelvins, muchísimo menos que la temperatura del espacio exterior.

⁸⁹ Es cierto que la empresa canadiense D-Wave ha fabricado dispositivos con varios cientos de qubits, pero se trata de ordenadores enfocados a resolver ciertas tareas de optimización concretas, no capaces de ejecutar todo tipo de algoritmos cuánticos.

⁹⁰ En inglés, *Noisy Intermediate-Scale Quantum technology* o *NISQ*.

tareas en las que más destacan. Por ejemplo, podrían emplearse al estilo de las actuales aceleradoras gráficas, recibiendo los datos y devolviendo las respuestas una vez realizados los cálculos. Eso sí, esa comunicación con los ordenadores tradicionales seguramente se llevará a cabo a través de la red, como una especie de *servicio en la nube*, dadas las especiales condiciones de refrigeración y aislamiento necesarias (ver Fig. 9).

En cualquier caso, disponer en el futuro de ordenadores cuánticos con los que resolver problemas que quedan más allá de las capacidades de la computación clásica supondrá, sin duda, un pequeño⁹¹ paso para las partículas involucradas, pero un gran paso para la humanidad.

Referencias:

- [1] Moore, G.E. “*Cramming more components onto integrated circuits*”. Electronics 38(8), 1965.
- [2] “*Is Moore’s Law still the law?*” Electronics Weekly, Sep 2017.
<https://www.electronicsworld.com/news/moores-law-still-law-2017-09/>
- [3] Wootters, W.K., Zurek, W.H. “*A single quantum cannot be cloned*”. Nature 299, 1982.
- [4] Nielsen, M.A., Chuang, I.L. *Quantum computation and quantum information: 10th anniversary edition*. Cambridge University Press, 2010.
- [5] Dawson, C.M., Nielsen, M.A. “*The Solovay-Kitaev algorithm*”. Quantum Information & Computation 6(1), 2006.
- [6] Deutsch, D., Jozsa, R. “*Rapid solutions of problems by quantum computation*”. Proceedings of the Royal Society of London A 439(1997), 1992.
- [7] Grover L.K. “*A fast quantum mechanical algorithm for database search*”. Proceedings, 28th Annual ACM Symposium on the Theory of Computing, 1996.
- [8] Dürr, C., Høyer, P. “*A quantum algorithm for finding the minimum*”.
<https://arxiv.org/abs/quant-ph/9607014>
- [9] Shor, P.W. “*Algorithms for quantum computation: discrete logarithms and factoring*”. Proceedings of the 35th Annual Symposium on Foundations of Computer Science, Santa Fe, 1994.
- [10] Rivest, R.L., Shamir, A., Adleman, L. “*A method for obtaining digital signatures and public-key cryptosystems*”. Communications of the ACM 21(2), 1978.
- [11] Feynman, R.P. “*Simulating physics with computers*”. International Journal of Theoretical Physics 21(6-7), 1982.
- [12] Harrow, A.W., Hassidim, A., Lloyd, S. “*Quantum algorithm for linear systems of equations*”. Physical Review Letters 103, 209.
- [13] Bennett, C.H., Brassard, G. “*Quantum cryptography: public key distribution and coin tossing*”. Proceedings of IEEE International Conference on Computers, Systems and Signal Processing, vol. 175, Nueva York, 1984
- [14] IBM Q Experience <https://quantumexperience.ng.bluemix.net/>
- [15] Preskill, J. “*Quantum computing in the NISQ era and beyond*”. Quantum 2(79), 2018.
- [16] Bauer, B., Wecker, D., Millis, A.J., Hastings, M.B., Troyer, M. “*Hybrid quantum-classical approach to correlated materials*”. Physical Review X 6, 2016.

Elías Fernández-Combarro Álvarez

Licenciado en Matemáticas e Ingeniero Técnico en Informática.

Doctor en Matemáticas.

Profesor Titular del Departamento de Informática, Área de Ciencia de la Computación e Inteligencia Artificial, Universidad de Oviedo.

Grupo de Computación Cuántica y de Altas Prestaciones.

⁹¹ Microscópico, literalmente.

Capítulo 22

Nuclear: la energía del futuro.

Es el año 2050, la mayoría de las personas del mundo viven en grandes ciudades. Coches totalmente eléctricos y autónomos circulan por las calles, casas inteligentes, y este verano te esperan tus primeras vacaciones en el espacio. Este es el futuro parece sacado de una película de ciencia ficción, pero como sabes, la realidad supera a la ficción. Solo nos separan 31 años de este escenario futurista que muchos imaginamos y esperamos. No son ideas disparatadas, como dijo DeGrasse Tyson hace solo 20 años que pensábamos que las nubes se encontraban únicamente en el cielo y ahora todo el conocimiento de la humanidad, conocido hasta la fecha, es albergado en una nube digital.

El progreso y la tecnología están en constante evolución y no se van a detener. Las demandas energéticas de semejantes avances han de ser satisfechas con las más innovadoras fuentes de energía. Entre las que se encuentra la nueva generación de energía nuclear, los nuevos diseños de reactores de **fisión nuclear** que darán tiempo para poder desarrollar los futuros reactores de **fusión**.



Ventana a tu futuro. Cortesía de ©Disney.

Los nuevos diseños de reactores de fisión deben promover la disponibilidad de sistemas y suministro a corto y largo plazo, y el aprovechamiento de combustible para la producción de energía en todo el mundo.

La fusión, por su parte, ha sido denominada el 'santo grial' de la tecnología energética, teniendo un gran potencial para ayudar a hacer frente al

desafío energético. La fusión nuclear está siendo ampliamente estudiada en todo el mundo como la **energía del futuro**.

Estas tecnologías no serán las únicas del futuro, las renovables sin duda estarán presentes en el mercado energético, pero supondrán la única alternativa energética capaz de suministrar la energía necesaria, de la que el ser humano podrá disponer con seguridad y abundancia, de varios órdenes de magnitud superior a cualquier otra.

En este capítulo se pretende aportar al lector una visión general, tanto de las nuevas tecnologías de fisión como las futuras plantas de fusión.

Fisión Nuclear.

La fisión nuclear ocurre cuando un núcleo de un átomo pesado captura un neutrón y se divide en dos o más núcleos de átomos más ligeros (productos de fisión), emitiendo grandes cantidades de energía. El calor generado mediante la fisión genera vapor que moverá la turbina para producir el movimiento de alternadores que transforman dicho trabajo mecánico en energía eléctrica en una central nuclear.

En este momento, gran parte del parque nuclear de todo el mundo va a desmantelarse y es preciso establecer un programa de nuevas instalaciones, las centrales del futuro. Estas centrales suponen un salto tecnológico y ya existen nuevos reactores instalados bajo el concepto de Generación III/III+, aunque otro grupo de reactores están en desarrollo, denominados Generación IV. Para estos últimos es necesario invertir tiempo y recursos en investigación de los diseños.

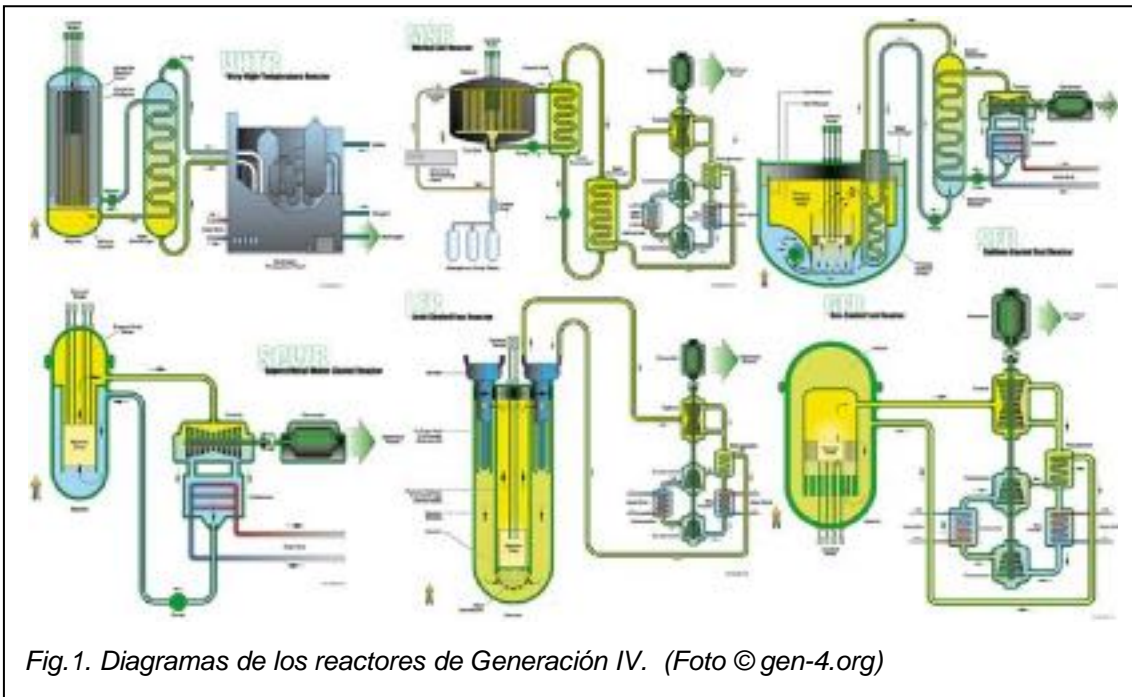
Estos reactores se basan en una serie de nuevos principios entre los que se encuentran: la sostenibilidad, economía, seguridad y fiabilidad. Los diseños destacan por su seguridad y fiabilidad, reduciendo al mínimo la probabilidad de daños en el núcleo del reactor y eliminar la necesidad de adoptar medidas de emergencia de fuera de la planta.

Las centrales de Generación III y III+ son una evolución basada en la experiencia adquirida de las centrales previas de Generación II. Estas mejoras afectan sobre todo a los sistemas de seguridad, operatividad y fiabilidad de las centrales y su estandarización. Las medidas de seguridad implementadas son sistemas de seguridad pasivos, cuyo funcionamiento se debe a fenómenos físicos como la gravedad sin necesidad de ningún tipo de energía eléctrica externa, y que actúan solos cuando la central se desvía de su modo normal de operación. Estas centrales ya están en construcción o funcionando en numerosos emplazamientos.

Las centrales de Generación IV, a diferencia de las centrales de Generación III, engloban nuevos diseños, pero sin olvidar las lecciones aprendidas de las centrales previas. Estos reactores utilizarán otro tipo de refrigerante a parte de agua, como el sodio o el helio, y serán capaces de utilizar como combustible el combustible usado en las plantas que están en

funcionamiento actualmente, lo que se denomina ciclo cerrado de combustible para la conversión eficiente de los residuos. No solo se podrán utilizar en generación eléctrica, sino que también se podrán aplicar en otros muchos campos como en la generación de hidrógeno, sistemas de transporte, exploración espacial o para generación térmica. En la Fig. 1 se pueden observar los esquemas y diagramas de las nuevas plantas de fisión.

Estos diseños no están en desarrollo actualmente ya que plantean grandes desafíos que están en fase de investigación sobre todo en materiales y combustible. Algunos de estos diseños de reactores podrían ver la luz en la próxima década, con despliegue comercial a partir de 2030. Tanto Francia como Rusia están desarrollando diseños avanzados de reactores rápidos de sodio para demostraciones a corto plazo. También se espera que se construya un prototipo de reactor rápido en Rusia en el marco de tiempo de 2020.



Existen también otro tipo de reactores nucleares denominados reactores modulares pequeños. En la Fig. 2 se puede observar un prototipo de estos nuevos reactores, en los que el reactor, contención y sistemas están localizados en un mismo módulo, lo que proporciona mayor seguridad ante posibles incidentes. Estos reactores se abrirán camino en el mercado para satisfacer la necesidad energética en un futuro a corto plazo y serán capaces de generar electricidad de forma flexible y asequible a todos los rincones del planeta para una variedad más amplia de usuarios y aplicaciones. Estos reactores modulares pueden desplegarse como central de un solo módulo o de varios módulos, ofreciendo la posibilidad de combinar la energía nuclear con fuentes de energía alternativas, incluida la energía renovable.



*Fig.2. Ejemplo de Reactor Pequeño modular NuScalePower Reactor.
(Foto © NuScalePower, LLC)*

Fusión Nuclear.

La fusión nuclear es la energía de las estrellas, la energía de nuestro sol. En su núcleo se producen reacciones de fusión de núcleos de hidrógeno para formar helio, liberando en el proceso grandes cantidades de energía electromagnética que llegan a nuestro planeta en forma de luz y calor.

Hace más de 50 años comenzó la investigación en fusión nuclear, pero todavía no se ha conseguido desarrollar un método capaz de generar fusiones nucleares en una tasa lo suficientemente alta para que la energía producida pueda ser utilizada para generación eléctrica con fines civiles.

La fusión nuclear es la reacción nuclear por la que dos núcleos de átomos ligeros, en general hidrógeno o sus isótopos (deuterio y tritio) se unen para formar uno de mayor peso atómico, convirtiéndose parte de la masa de los reactivos en energía.

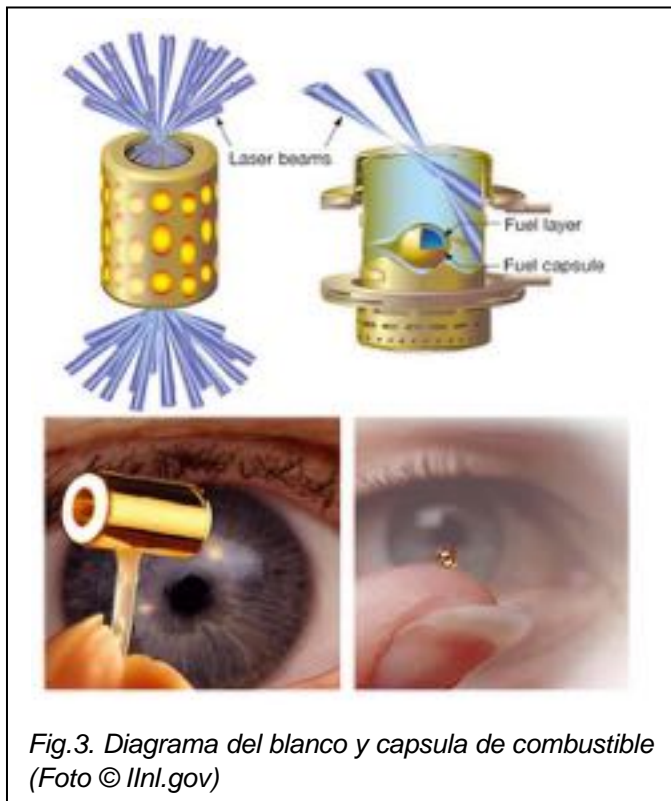
La opción de combustible más utilizada y la más viable hasta la fecha es la reacción de deuterio y tritio. El deuterio es un isótopo del hidrógeno que se encuentra de manera natural, pudiéndose extraer de la molécula de agua, por lo que los océanos serían una fuente casi inagotable de este producto. El tritio por su parte es un isótopo artificial y por tanto debe ser producido a partir de irradiación neutrónica de litio.

Para obtener reacciones de fusión se debe crear un medio tan denso que los núcleos no tengan casi ninguna posibilidad de escapar, superando las fuerzas de repulsión, y se fusionen entre sí. La primera de las dificultades encontradas para su desarrollo es que el proceso se debe iniciar mediante un gran aporte energético. Otras de las dificultades para producir reacciones de fusión en una tasa elevada, es que el combustible ha de ser confinado, comprimido y calentado hasta alcanzar un estado de plasma, compuesto por electrones libres y átomos ionizados, con temperaturas de millones de grados y presiones del orden de Mbar (un millón de veces la presión atmosférica). Las

condiciones para la obtención de la fusión nuclear se resumen en una combinación adecuada de temperaturas, densidades y confinamiento que se denominan criterio de Lawson.

Un plasma siempre tiende a expandirse libremente como lo hace un gas, que tiende a ocupar la totalidad del recipiente donde está contenido. En la cavidad de un reactor de fusión nuclear es necesario garantizar el confinamiento y control del plasma el tiempo necesario para que se produzca la reacción, para ello es necesario lograr una densidad del plasma para que los núcleos estén cerca y puedan dar lugar a las reacciones de fusión. Sin embargo, los confinamientos convencionales, con paredes de vasija, no son factibles debido a las altas temperaturas del plasma. Por lo tanto, para confinar el plasma se necesita invertir energía en mantenerlo en unas determinadas condiciones de presión, densidad y temperatura. Hay dos métodos posibles de confinamiento del plasma: fusión por confinamiento inercial, y fusión por confinamiento magnético. En el primero tipo, una diminuta cápsula de combustible se comprime mediante láseres de gran potencia, de manera que el plasma se confina durante muy poco tiempo, pero a densidades extremadamente altas. En el confinamiento magnético, un plasma de baja densidad se confina mediante campos magnéticos que lo mantienen en trayectorias cerradas. En este tipo de confinamiento existen dos diseños: los Tokamak, con forma toroidal, y los Stellarators, con forma helicoidal más compleja.

Vamos a explicar con más detalle cada uno de los tipos de confinamiento, las posibles centrales del futuro y los retos a los que se tendrá que hacer frente para hacerlas posible.



Fusión por Confinamiento Inercial (FCI).

En el proceso de ignición en la fusión por confinamiento inercial, una cápsula de combustible (deuterio y tritio), se comprime por la aceleración que sufre debido a la absorción de gran cantidad de energía, provocando que su parte exterior se evapore y que, por conservación del momento cinético, la parte interior se vea empujada hacia dentro, comprimiéndose y calentándose.

La cápsula de combustible se denomina “blanco de fusión” con un

diámetro de unos milímetros, que contiene en su interior el combustible. La parte exterior de la cápsula se denomina ablator y la interior pusher debido a la función que cumplen cada uno en el proceso. En la Fig. 3 se puede ver el diagrama de la cápsula y donde está contenida, así como el tamaño de estas en comparación con la pupila humana.



Fig.4. Cámara de blancos en las instalaciones de NIF. (Foto ©Inl.gov)

La energía necesaria para la ignición se deposita mediante una fuente denominada sistema accionador o driver, que consigue que la cápsula pase a estado de plasma y parte del combustible comience a fusionar. En este caso, este sistema accionador se compone por un sistema de láseres. Por ejemplo, en la instalación de prácticas NIF (National Ignition Facility) en EE. UU, el sistema de láseres consta de 192 haces de láser, capaces de generar 2.1 Mega Julios de potencia, récord batido en julio de 2018.

Estos láseres hacen blanco en la cápsula de combustible en una cámara, denominada cámara de blancos, de forma esférica con 10 metros de altura. En la Fig. 4 se puede ver el exterior de dicha cámara con el sistema de láseres apuntando hacia el interior de esta. En la siguiente figura, Fig. 5, **¡Error! No se encuentra el origen de la referencia.** se tiene una visión del interior de la cámara de blancos, donde se observan los agujeros de los 192 láseres que apuntan hacia el centro de esta.

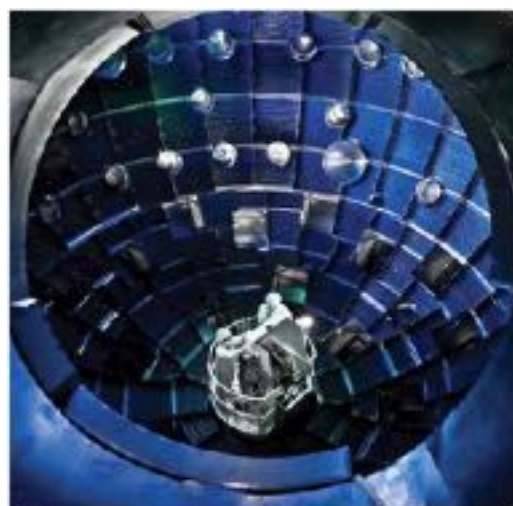


Fig.5. Cámara donde se sitúan los blancos en NIF. (Foto ©Inl.gov)

El objetivo a largo plazo es que en una hipotética central de fusión inercial se produciría un ritmo constante de igniciones de cápsulas en la cámara de blancos. Los neutrones calentarían la cámara y esta energía a su vez se transferiría a un refrigerante situado en las paredes, siendo un ciclo de producción de energía similar al de cualquier central de producción eléctrica actual que opera con una turbina.

Actualmente, no está disponible ningún ejemplo de planta de potencia para uso civil o comercial, solo se encuentran instalaciones de investigación. La instalación de prueba más avanzada de un reactor de fusión por confinamiento inercial es el National Ignition Facility (NIF), construida en el Laboratorio Nacional Lawrence Livermore (LLNL) en EE. UU y Laser MegaJoule en Francia operativa desde octubre 2018. El NIF está diseñado para la demostración de producción de energía de fusión con un disparo simple, es decir, demostración de ganancia de energía con un solo blanco. Las figuras, Fig. 3-5, muestran la instalación experimental NIF. La tarea principal de la instalación Laser MegaJoule son los cálculos de refinamiento para la fusión por confinamiento inercial. La instalación consta de 240 tubos de láseres capaces de generar 1,8 MJ. En la actualidad, se ha terminado la fase de construcción de la instalación que tendrá fines militares únicamente, aunque se dejará la realización de algún experimento civil (Fig. 6), la instalación se localiza en el sur de Francia en un emplazamiento secreto.



La prolongación natural de un proyecto como el NIF sería que las condiciones de un solo disparo, el uso de un único blanco, se produjeran de forma repetitiva tanto en el sistema de láser como el de inyección de blanco. El objetivo de dicha instalación sería la capacidad de producir energía y producir tritio como parte de un ciclo cerrado de producción de combustible y poder tener todo lo necesario para una planta de fusión operativa en el mismo emplazamiento. Este nuevo proyecto, también desarrollado por el laboratorio LLNL, se denomina Inertial Fusion Engine (LIFE). Se ha desarrollado un diseño de sistemas para LIFE, ver boceto de la Fig. 7 para la planta LIFE, que podría

llevar a una central eléctrica operativa a mediados de la década del 2020 y distribución de energía generada por fusión en la década de 2030. Aunque todavía queda mucho trabajo de desarrollo técnico por delante.

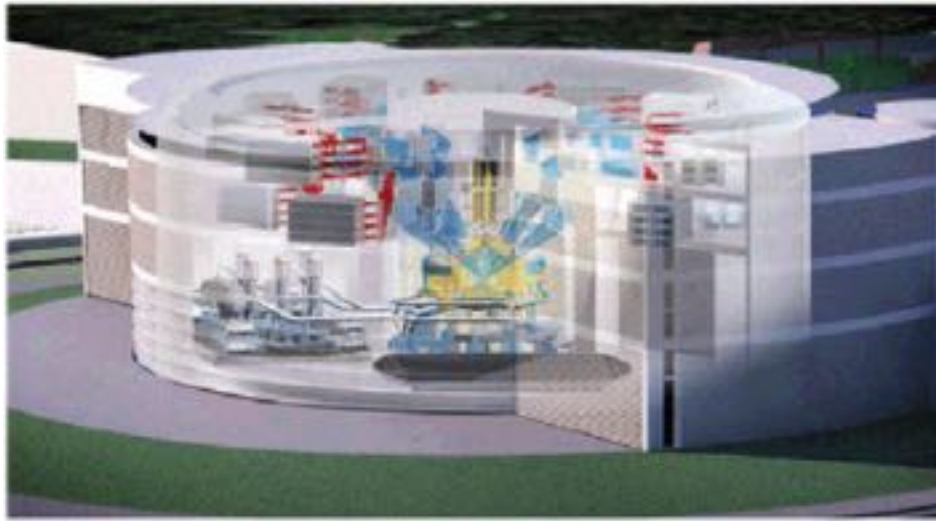


Fig.7. Esquema de lo que será la futura planta de fusión con el proyecto LIFE (Foto © LLNL)

Al otro lado del océano se encuentra la instalación HiPER (High Power laser Energy Research, Investigación sobre Energía con Láseres de Alta Potencia) que permitirá a Europa tomar el liderazgo en cuanto al desarrollo de la obtención masiva de energía por fusión Inercial. Este proyecto tiene como principal objetivo un enfoque para la producción de energía a nivel de planta comercial. Para ello se ha de pasar de la demostración de laboratorio con un sistema más eficiente de compresión y calentamiento a soluciones tecnológicas más avanzadas en materiales y, al igual que LIFE, con un sistema repetitivo de blancos. Este nuevo enfoque establecerá una línea temporal a largo plazo para la generación de energía en condiciones de reactor o planta de potencia de uso civil. En la Fig. 8 se puede ver el diseño de la planta de reactor HiPER, la instalación se encuentra en la fase de diseño y evaluación de tecnología con una fecha de comienzo de construcción no antes del 2025.

Las mayores diferencias entre el proyecto LIFE y el proyecto HiPER son en el tipo de blanco utilizado y la protección interna de las paredes del reactor, la cámara de blancos.

LIFE utilizará los denominados blancos indirectos como los de NIF. En estos blancos, la esfera de combustible esta introducida en una cápsula (ver Fig. 3). La energía láser no se deposita directamente sobre la esfera de combustible, sino que primero se convierten en rayos X en las paredes internas de una cápsula. Estos rayos X comprimen y calientan el combustible hasta su ignición. Este tipo de blancos son menos eficientes, pero están más desarrollados y tienen menos inestabilidades a la hora de la compresión del combustible. HiPER utiliza los denominados blancos directos, en los que la energía del láser se deposita directamente en la esfera de combustible.

La segunda diferencia está en el revestimiento del reactor. Esta diferencia nace también de la diferencia de blancos utilizados. En ambos blancos el 70% de la energía de fusión es emitida en forma de neutrones energéticos. Mientras que en los blancos indirectos (LIFE) el 30% restante es fundamentalmente radiación X, en los blancos directos (HiPER) ese 30% son iones rápidos. Esta diferencia hace que se utilicen diferentes modelos de paredes de reactor en ambos casos: en el caso de LIFE, para contrarrestar el efecto de los rayos X, la cámara se llenará de gas Xe que absorberá la energía y la depositará sobre las paredes; y en el caso de HiPER, la pared del reactor constará de una pared interna de wolframio que hará de escudo protector.

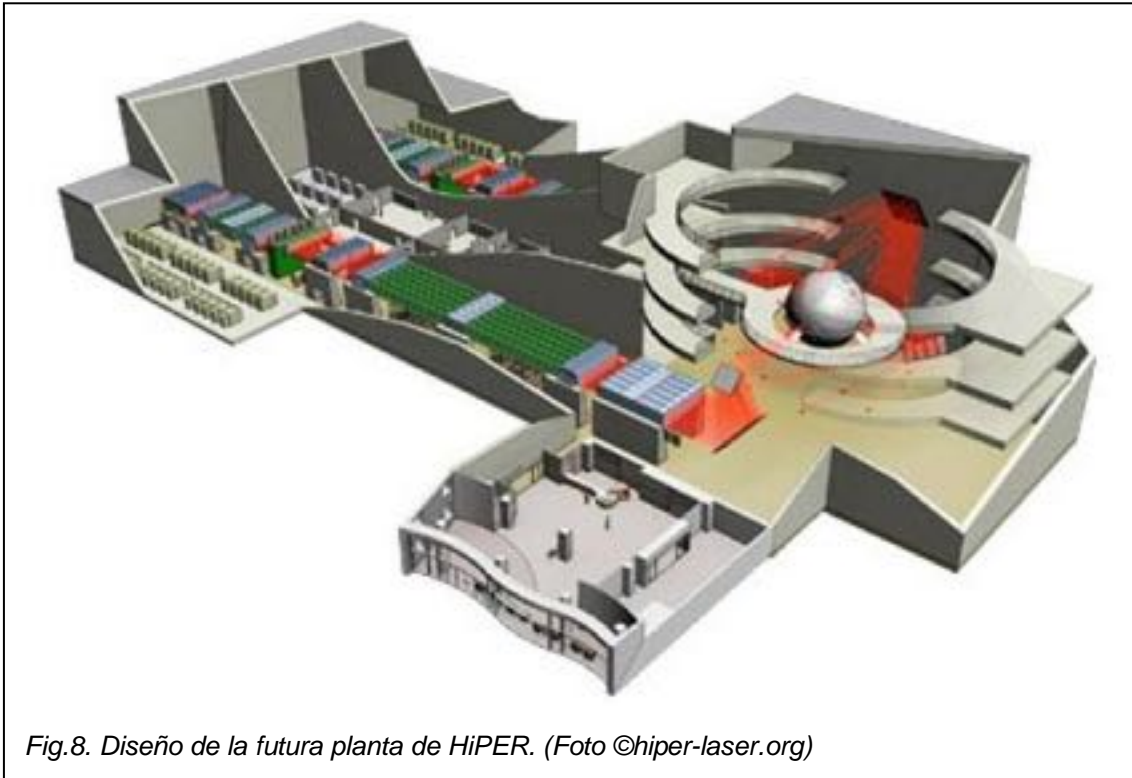


Fig.8. Diseño de la futura planta de HiPER. (Foto ©hiper-laser.org)

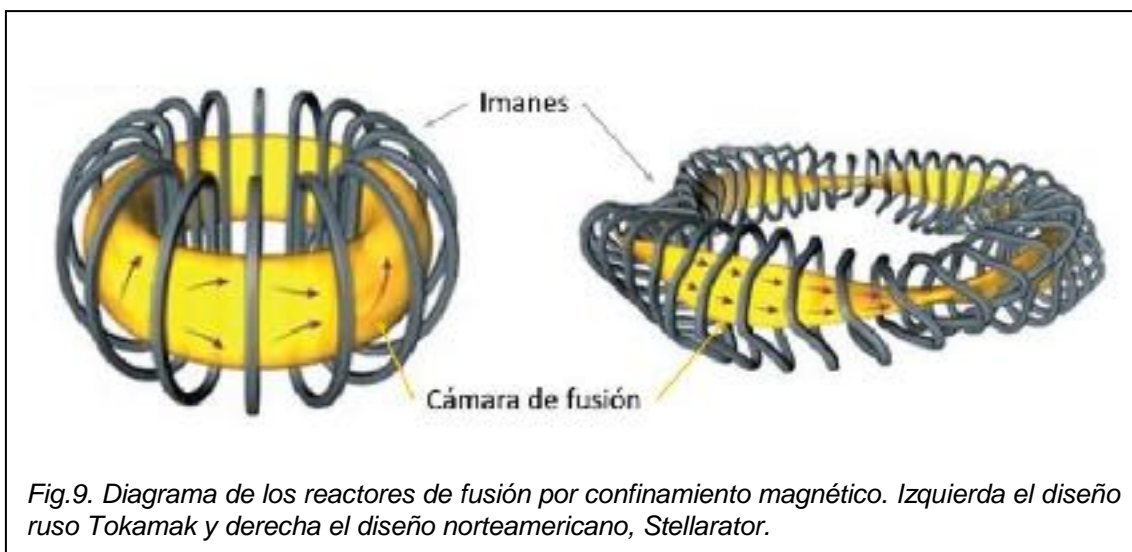
Fusión por Confinamiento Magnético (FCM).

Como se ha descrito en la introducción, para alcanzar la ignición se requiere producir y mantener un plasma confinado durante el tiempo suficiente y bajo las condiciones necesarias como para que se produzca un número suficiente de reacciones de fusión. Una de las opciones para conseguirlo es el denominado confinamiento magnético.

En el confinamiento magnético, el combustible es calentado de manera que los choques entre partículas son producidos por agitación térmica. Las partículas eléctricamente cargadas del plasma son controladas mediante campos magnéticos. A lo largo de estas líneas de campo magnético, las partículas deben moverse describiendo hélices. Las líneas se cierran sobre sí mismas, contenidas en una región limitada de espacio, las partículas quedarán confinadas en densidades no tan elevadas, pero durante más tiempo por lo que se podrán conseguir más reacciones de fusión. La fuerza de Lorentz es la

encargada de desviar la trayectoria de las partículas del plasma y mantenerlas en el sistema. Debido a las fuerzas electromagnéticas implicadas, no es suficiente el campo toroidal para confinar el plasma, y se hace preciso modificar éste con una componente poloidal. Es en el método seguido para generar esta componente donde reside la principal diferencia entre las dos tecnologías principales: el Tokamak y el Stellarator.

En la Fig. 9 se muestran un esquema de las dos tecnologías, en la izquierda el Tokamak y en la derecha el Stellarator, donde se observan a simple vista las diferencias entre ambos sistemas. Los primeros modelos de fusión magnética desarrollados fueron los norteamericanos Stellarator con un campo generado directamente en un reactor toroidal, con el problema de que el plasma se filtraba entre las líneas del campo. El dispositivo más desarrollado para fusión por confinamiento magnético es el diseño Tokamak. Sin embargo, el mayor reactor de este tipo, el JET (Joint European Torus, 'Toro Común Europeo') no ha logrado mantener una mezcla a la temperatura (1 millón de grados) y presiones necesarias para que se mantuviera la reacción a largo plazo. Situado en UK, JET es el resultado del consorcio de más de 40 laboratorios europeos. Se estima que estará operativo hasta finales de 2019. Diseñado expresamente para estudiar el comportamiento del plasma en condiciones cercanas a las necesarias en un reactor. Produjo su primer plasma en 1983, y su mayor hito fue cuando consiguió la fusión nuclear controlada en 1997 (17 MW en un segundo). Este reactor ha sido modificado, como se puede ver en la Fig. 10, para asemejarse al que será el próximo gran proyecto de fusión por confinamiento magnético, el ITER, explicado a continuación.



Desde 1986, en un esfuerzo internacional para lograr la fusión, se ha comprometido la creación del mayor reactor de fusión tipo Tokamak, denominado ITER (Reactor Termonuclear Experimental Internacional), proyecto liderado por la Unión Europea y en el que participan 35 naciones incluida EE.UU., Rusia, Japón, India, Corea del Sur y Suiza.

En la actualidad ya se han alcanzado las temperaturas y densidades necesarias para producir fusión en los tokamaks existentes, pero la ganancia

de energía está limitada por el tamaño de las instalaciones. A mayor envergadura, mayor será la ganancia de energía y ese es el objetivo de ITER, el cual tendrá dos veces el tamaño de JET (el Tokamak con mayor tamaño hasta ITER). ITER se está construyendo desde el 2010 en el sur de Francia en Cadarache y se espera que obtenga su primer plasma en 2035.



Fig.10. Vista interior del Tokamak europeo JET, donde se está mejorado la pared y el desviador para asemejarlos a ITER. (Foto © iter.org)

El objetivo de ITER es producir unos 500 MW, usando 50 MW para la ignición de un gramo de combustible de tritio durante al menos unos 500 s. En la Fig. 11 se muestra un diagrama del reactor ITER. Este reactor sigue siendo un reactor experimental, es decir, no está destinado a producir energía de modo continuo para su distribución a la red eléctrica.

Para lograr la generación de energía se necesitaría construir otro prototipo de reactor comercial. Este proyecto se llama DEMO (DEMONstration Power Plant). Este proyecto tiene como fin la producción de energía en modo continuo, se espera que pueda producir 1.5GW de energía eléctrica (un reactor de fisión actual genera entorno a los 900MW). Su construcción está prevista que comience alrededor de 2030, para estar en operación en 2040 y cumplir sus objetivos antes del 2050. El diseño de DEMO no ha sido seleccionado formalmente, se basará en las lecciones aprendidas en cuanto a física y tecnología validada en ITER anteriormente. Los principales objetivos de DEMO serían: la producción de energía con conexión a la red eléctrica, obtener la cantidad de tritio necesaria para cerrar el ciclo de combustible y demostrar toda la tecnología para la construcción de futuras centrales comerciales.

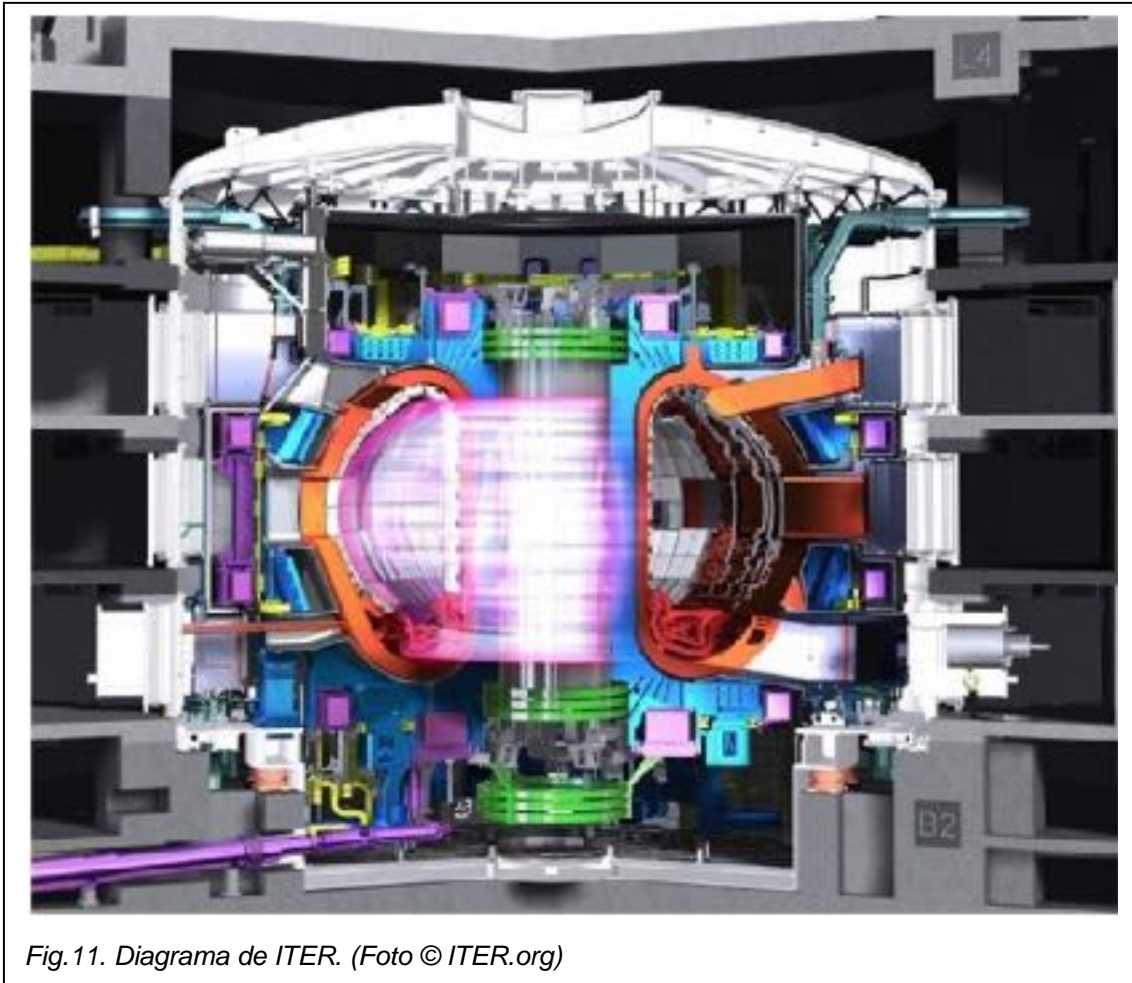


Fig.11. Diagrama de ITER. (Foto © ITER.org)

Una vez sentadas las bases para futuros reactores comerciales se necesitarán diseños para esos reactores de futuro, ahí entra PROTO (PROTOTYPE Power Plant), el primer prototipo de reactor comercial de fusión comercial, aunque no está diseñado en detalle, la propuesta se basa en simultanear la construcción de un proyecto combinado DEMO/PROTO con ITER. Se introducirían mejoras en ITER y se diseñaría DEMO de tal forma que permitiera su transformación a PROTO sin necesidad de otro proyecto por separado

Retos por cumplir.

Para demostrar la viabilidad de la fusión, tanto por confinamiento inercial como por confinamiento magnético es necesario todavía un gran desarrollo de la tecnología e ingeniería relacionada con materiales, códigos, diseños, etc. Actualmente existen numerosos proyectos en marcha con este propósito, como por ejemplo el caso de IFMIF (International Fusion Material Irradiation Facility). IFMIF forma parte de los programas internacionales para la consecución de la fusión como fuente de energía y tiene como objetivo el reproducir las mismas condiciones de radiación duras del interior de un reactor de fusión por confinamiento magnético para la validación y estudio de materiales para un

futuro reactor de fusión. Esta planta se considera clave para poder diseñar DEMO en base no solo a la información obtenida en ITER sino también en base la información obtenida sobre el comportamiento de los materiales bajo la radiación neutrónica esperada.

Las instalaciones experimentales, tanto ITER como NIF, permitirán además de demostrar que es posible dominar la física de la fusión nuclear en los dos caminos termonucleares posibles, disponer de un banco de pruebas para el desarrollo de las nuevas tecnologías que sustenten un reactor de fusión a nivel comercial. Se necesita también ir dando los pasos necesarios para diseñar y demostrar los sistemas que compondrán dichos reactores de potencia.

Algunos de los retos a los que hay que hacer frente para lograr las plantas de potencia, tanto inercial como magnética, están mencionados brevemente a continuación:

- Tecnología eficiente de los superconductores que generarán los campos magnéticos deseados en la FCM.
- La consecución de largos tiempos de confinamiento que lleven a un plasma con suficientes reacciones de fusión.
- Obtener unos blancos de alta ganancia con energía generada por FCI del orden de 50 a 1000 veces mayor que aquella proporcionada por los láseres.
- Los materiales capaces de aguantar las cargas térmicas y las altas dosis de radiación.
- Diseño y elección final del manto reproductor de tritio para poder regenerar tritio que alcance no solo a reponer lo gastado, sino producir al menos un 20% más.
- Fabricación de blancos de bajo coste.
- Cámara de reacción de larga vida operativa.
- Definición y establecimiento de los niveles de radiactividad en los diversos puntos del sistema.

Solo destacar que algunos de los objetivos en el área de tecnología del reactor de potencia como la cuestión de materiales, tritio o sistemas de planta son comunes a ambas tecnologías de fusión.

Punto final.

¿Merece la pena?, ¿merece la pena todo el esfuerzo e inversión en estas nuevas tecnologías? La contestación a esta pregunta es sencilla, SI. Merece la pena toda la investigación que se está llevando a cabo para poder resolver el problema que amenaza las sociedades del siglo XXI y posteriores, la falta de suministro eléctrico.

Los nuevos reactores de fisión y las futuras plantas de fusión nuclear, en una u otra de sus versiones tecnológicas, son la opción más clara para el futuro del abastecimiento energético de una sociedad que cada vez demanda y demandará más energía.

Sin olvidarnos claro está de otro tipo de energías, como las energías renovables que tendrán un papel fundamental en nuestro futuro energético esperando un incremento de su uso que pasará de un 25% actualmente a 37% en 2040.

Se debe asumir un modelo de generación eléctrica que debe combinar diferentes fuentes de energía sostenibles con el medio ambiente (nuclear y renovable) con una óptima red de distribución. Así se asegurará el suministro de energía necesario para ver el futuro que todos esperamos.

Referencias:

- [1] Jóvenes Nucleares. Sociedad Nuclear Española. Curso básico de Fusión Nuclear. Septiembre de 2017, Madrid. ISBN: 978-84-697-5718-5
- [2] Jóvenes Nucleares. Sociedad Nuclear Española. Curso Básico de Ciencia y Tecnología Nuclear. Depósito legal: M-24599-2013
- [3] Advances in Small Modular Reactor Technology Developments. IAEA Advanced Reactors Information System (ARIS). IAEA. 2018 Edition
- [4] Reactores de fisión avanzados: <http://www.world-nuclear.org/information-library/nuclear-fuel-cycle/nuclear-power-reactors/advanced-nuclear-power-reactors.aspx>
- [5] Reactores de fisión de Generación IV: <http://www.world-nuclear.org/information-library/nuclear-fuel-cycle/nuclear-power-reactors/generation-iv-nuclear-reactors.aspx>, https://www.gen-4.org/gif/jcms/c_9260/public
- [6] Fusión nuclear: <https://energia-nuclear.net/que-es-la-energia-nuclear/fusion-nuclear>
- [7] ITER: <https://www.iter.org/>
- [8] NIF: <https://lasers.llnl.gov/>
- [9] HiPER: <http://www.hiper-laser.org/>
- [10] Laser MegaJoule: <http://www-lmj.cea.fr/index-en.htm>
- [11] IFMIF: <http://www.fusion.ciemat.es/Inf-en-proyectos-internacionales/ifmif/>

Emma López-Alonso Conty
Doctora en Ciencia y Tecnología Nuclear.

Capítulo 23

Nanotecnología para la salud: cómo lo muy pequeño ayuda en medicina.

Hace unos dos mil cuatrocientos años, Demócrito desarrolló la primera teoría atómica para explicar la realidad en que vivimos. Pensó que toda la materia estaba formada por una mezcla de elementos inmutables y eternos, unas entidades infinitamente pequeñas, indivisibles e imperceptibles por los sentidos, a las que denominó átomos. Resulta asombroso pensar que Demócrito llegó mediante simples razonamientos lógicos a esa descripción bastante acertada del mundo a escala nanométrica (recordemos que un nanómetro es la milmillonésima parte de un metro). Por supuesto, hoy en día sabemos que el átomo sí es divisible en un núcleo formado por protones y neutrones y una nube electrónica alrededor, pero las entidades más pequeñas que podemos manipular para fabricar objetos son los átomos. De hecho, podemos definir la nanotecnología como la manipulación de átomos y moléculas en la nanoescala para obtener estructuras o dispositivos que, gracias al control de su forma, tamaño y composición, sirven para cumplir una determinada función.

Introducción: herramientas en nanotecnología.

Han sido los avances instrumentales durante el siglo XX quienes han permitido el desarrollo de la nanotecnología, pues es necesario visualizar esos objetos en la nanoescala, comprender los fenómenos a que dan lugar, y también fabricarlos de forma controlada para optimizar la propiedad deseada.

De entrada, como el tamaño de los átomos es unas mil veces inferior a la longitud de onda de la luz visible (400-700 nm), no pueden observarse utilizando instrumentos ópticos convencionales, por lo que hubo que esperar a la invención de los microscopios electrónicos y los microscopios de sonda de barrido para poder visualizar esos objetos en la nanoescala. Entre los primeros, cabe distinguir entre los de transmisión (TEM, por las siglas en inglés de *Transmission Electron Microscope*), que emplean electrones que atraviesan la muestra a estudio con un voltaje de aceleración entre 100 y 300 kV, y los de barrido (SEM, *Scanning Electron Microscope*), que trabajan con los electrones reflejados y/o re-emitidos, con voltajes típicos entre 1 y 30 kV. Estos microscopios permiten estudiar no solo la morfología de los materiales en la nanoescala, sino además su composición química y su estructura cristalina. En cuanto a los microscopios de sonda de barrido (SPM, *Scanning Probe Microscope*), consisten en aproximar mucho, a distancia nanométrica, una sonda a la muestra a estudio, y medir entonces una determinada magnitud. El

primero en desarrollarse fue el microscopio de efecto túnel (STM, *Scanning Tunneling Microscope*), que se basa en un efecto cuántico: la corriente túnel que aparece entre una punta metálica y una muestra conductora cuando se establece una diferencia de potencial entre ellas. Posteriormente, se desarrolló el microscopio de fuerzas atómicas (AFM, *Atomic Force Microscope*), que ya podía emplearse con muestras no conductoras, y luego distintas variantes (el microscopio de fuerzas magnéticas, el de fuerzas electrostáticas, el de potencial superficial, el de fricción, el de campo cercano, etc.) que permiten analizar diversas propiedades: magnéticas, eléctricas, mecánicas, ópticas...

En paralelo, se han desarrollado las técnicas que permiten la fabricación de nanoestructuras y nanodispositivos, que puede llevarse a cabo siguiendo dos aproximaciones, denominadas descendente (*top-down*) y ascendente (*bottom-up*), o también una combinación de ellas, y en ese caso se habla de una ruta convergente. La vía descendente consiste en miniaturizar o moldear desde lo macroscópico. Las técnicas más empleadas en esta ruta son la molienda, el ataque por iones, y diversos tipos de litografía: por electrones, por haz de iones, por luz ultravioleta (cuanto menor es la longitud de onda más pequeños son los motivos que se pueden diseñar: en la actualidad ya se emplea el ultravioleta extremo) o por nanoimpresión. Por su parte, en la vía ascendente se ensamblan nanoentidades (átomos, moléculas), tanto de forma forzada como mediante procesos de autoensamblado. Las técnicas más usuales en esta aproximación son la síntesis química, la síntesis electroquímica, la manipulación por SPM, los procesos sol-gel, y las diversas técnicas de condensación desde la fase vapor: la deposición química (CVD, *Chemical Vapor Deposition*), la deposición de capas atómicas (ALD, *Atomic Layer Deposition*), la epitaxia de haces moleculares (MBE, *Molecular Beam Epitaxy*), la ablación láser (PLD, *Pulsed Laser Deposition*), y la pulverización catódica (*sputtering*). Estas últimas técnicas pueden dar lugar además a fuentes de nanopartículas cuando se emplea una cámara de agregación [1].

Asimismo, en los últimos años ha cobrado importancia la síntesis biológica o biosíntesis para la producción de nanopartículas. Es una rama basada en la química verde (se denomina así a la que usa materiales naturales y procesos sostenibles) que se dedica a emplear entidades biológicas unicelulares y multicelulares (como bacterias, hongos, plantas, virus, levadura...) que actúan como pequeñas factorías para sintetizar nanopartículas con una amplia gama de tamaños, formas, composiciones y propiedades fisicoquímicas [2]. Es un método limpio, no tóxico y respetuoso con el medio ambiente, y que debido a su bajo costo puede utilizarse en países en vías de desarrollo.

En definitiva, poseemos las herramientas que nos permiten fabricar y caracterizar un amplio conjunto de materiales con una, dos o tres de sus dimensiones de tamaño nanométrico. Por ejemplo, los discos duros de ordenador y los paneles fotovoltaicos están formados por multicapas, es decir, un apilamiento de láminas delgadas, cada una de ellas de composición uniforme y con una dimensión, su espesor, de tamaño nanométrico. Un caso ilustrativo con dos dimensiones en la nanoescala son las nanofibras de carbono, que debido a sus propiedades mecánicas se utilizan en la fabricación de raquetas de tenis, palos de golf y materiales para aeronáutica. Y con tres

dimensiones nanométricas podemos mencionar las nanopartículas de dióxido de titanio, que ya se emplean en protectores solares porque filtran los rayos ultravioleta.

Puede afirmarse, por tanto, que la nanotecnología ya está aquí y ha venido para quedarse. ¿Dónde podemos esperar que tenga un mayor impacto? Pues sobre todo en tres sectores de gran interés económico y social: tecnologías de la información y las comunicaciones, energía y medioambiente, y ciencias de la salud. En lo que sigue, nos centraremos en este campo, en el que los nanomateriales van a interactuar con entidades en la nanoescala de unos pocos nanómetros (moléculas orgánicas, sin ir más lejos el diámetro de la doble hélice del ADN es de unos 2 nm), de una decena de nm (anticuerpos), centenares de nm (virus y bacterias) e incluso de miles de nm (células). Analizaremos su aplicación en dos vertientes: en las técnicas de diagnóstico y en los tratamientos para curar enfermedades.

Nanotecnología para técnicas de diagnóstico.

Un primer ejemplo de cómo los nanomateriales ayudan en las técnicas de diagnóstico lo ofrecen las nanopartículas que mejoran las prestaciones de tecnologías de imagen médica. Estas se basan en que aparezcan contrastes en la forma en que la luz se dispersa o se emite en el tejido sano en comparación con el tejido enfermo. Cuanto más definido sea ese contraste, o mayor resolución tenga, más preciso será el diagnóstico. Ese es el caso de las nanopartículas magnéticas, que pueden emplearse como agentes de contraste en los equipos de imagen por resonancia magnética [3]. Esta técnica de imagen está basada en medir la resonancia magnética nuclear, en particular mediante la alteración del momento magnético de los protones del agua contenida en el cuerpo humano. Se comienza aplicando un intenso campo magnético externo continuo que alinea esos momentos, se les somete entonces a pulsos de campo alterno de radiofrecuencia, y, a través de la señal de radiofrecuencia que emiten, se mide cuánto tardan en recuperar su posición de equilibrio alineados al campo continuo: son los denominados tiempos de relajación. Cuando se emplean nanopartículas magnéticas (por ejemplo de óxidos de hierro), ellas crean un pequeño campo magnético local que provoca una disminución en los tiempos de relajación en su entorno, mejorándose por tanto el contraste allí donde están situadas. Por lo tanto, lo que hay que hacer previamente es funcionalizarlas adecuadamente, para que tras ser inyectadas por vía intravenosa, se unan al tejido o tipo de células que se desea estudiar.

Aparece aquí uno de los conceptos clave en nanotecnología: la funcionalización de los nanomateriales, es decir, dotarles de funciones o propiedades adicionales, generalmente modificando su superficie. Un caso típico es emplear un recubrimiento que les dote de las características adecuadas: que no se deterioren en el medio (p.ej. que resistan el pH, que tengan la hidrofobicidad o hidrofiliicidad necesaria, etc.), que lleven adherido el ligando indicado para unirse al tipo de tejido o de células deseado, o que porten el medicamento que se desea administrar localmente. Entre los diferentes recubrimientos y biomoléculas empleados para funcionalizar las

nanoestructuras podemos mencionar antígenos y anticuerpos (se unen específicamente entre sí), aptámeros (alta afinidad específica a determinadas moléculas), cadenas de ADN (para hibridarse con su cadena complementaria), polímeros, o incluso virus.

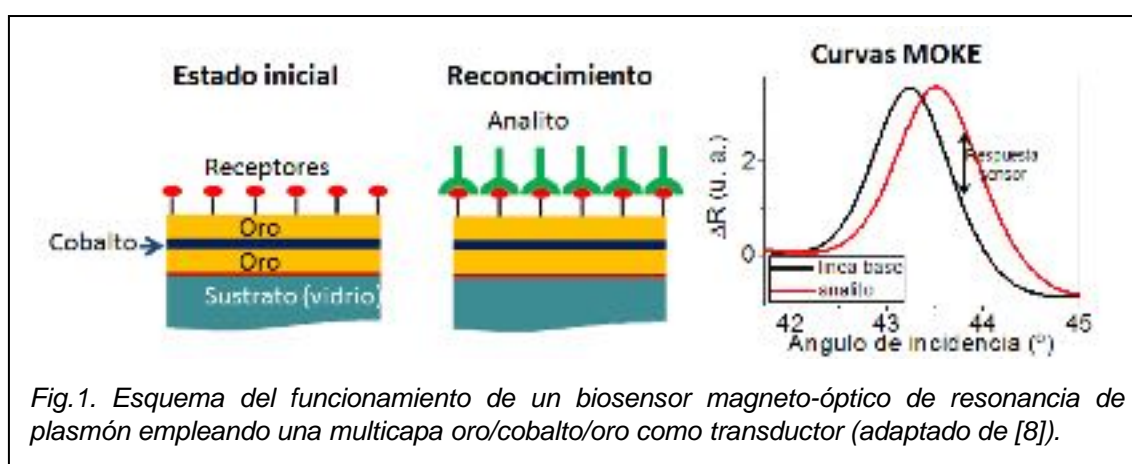
Una vez entendido lo que es la funcionalización, podemos seguir mencionando aplicaciones de las nanoestructuras en las técnicas de imagen. Otro ejemplo interesante lo proporcionan nanopartículas con iones de tierras raras, como Europio o Gadolinio, que presentan luminiscencia persistente en el rango espectral del infrarrojo cercano (NIR, del inglés *near-infrared*). Estas nanopartículas pueden ser excitadas ópticamente *in vitro* antes de ser inyectadas en un animal de laboratorio, y su distribución se puede seguir en tiempo real durante varias horas con un microscopio óptico sin necesidad de ninguna fuente de iluminación externa. Mediante la funcionalización adecuada, es posible por tanto emplearlas para imagen *in vivo* de tumores y reparto local de fármacos [4]. Además, empleando distintos iones en la misma nanopartícula o distinto dopado, es posible preparar partículas multimodales, es decir, que emitan en regiones distintas del NIR: la luminiscencia en la segunda ventana del NIR (longitud de onda entre 1000 y 1350 nm) tiene una relación señal/ruido más alta y una penetración en el tejido más profunda que luminiscencia en la primera ventana del NIR (650 a 950 nm), debido a una menor absorción y dispersión de fotones y a una mínima autoluminiscencia de los tejidos en aquel rango [5].

Por otro lado, los nanomateriales también pueden utilizarse en biosensores, que son dispositivos que detectan la presencia de determinadas especies biológicas, por ejemplo antígenos procedentes de virus o los anticuerpos que genera el sistema inmunológico para responder a aquellos. Los nanomateriales actúan como transductores en un biosensor, es decir, el componente que hace de detector en el que cambia una propiedad física o química cuantificable cuando se produce la interacción biológica, de unión o de reconocimiento, con el analito (la especie que se desea detectar). Para que eso suceda, la nanoestructura tiene que estar convenientemente funcionalizada con el elemento biológico que interacciona con esa especie en particular. La propiedad medible puede ser muy diversa: efectos ópticos, efectos piezoeléctricos, respuesta mecánica, efectos magneto-ópticos, respuesta magnética, etc.

Analicemos un caso paradigmático: los biosensores basados en resonancia de plasmón [6]. El nanomaterial es una lámina de espesor nanométrico de un metal noble, normalmente oro, y la resonancia de plasmón es una excitación colectiva de sus electrones de conducción que se propaga como onda evanescente, lo que se manifiesta con una absorción total de la luz en un ángulo particular de incidencia. Ese ángulo en el que ocurre la absorción total depende del índice de refracción del medio con el que el oro está en contacto. Para que actúe como detector, su superficie se funcionaliza con la especie complementaria al analito deseado: si por ejemplo se desea detectar un anticuerpo, sobre el oro se inmoviliza su antígeno específico. Si se produce el reconocimiento, es decir, cuando el anticuerpo se une al antígeno inmovilizado sobre el oro, la composición del medio con el que el oro está en contacto se modifica, por lo que se altera el índice de refracción y por

consiguiente la resonancia de plasmón se ve modificada, lo que se manifiesta en un cambio en la señal óptica.

Una mejora en este dispositivo es usar no solo una lámina de oro sino una multicapa en la que el oro se combina con otra lámina de espesor nanométrico de un material magnético, por ejemplo hierro o cobalto. El funcionamiento es similar, con la diferencia de que se emplean unas pequeñas bobinas o un imán para aplicar un campo magnético, y se mide entonces un efecto magneto-óptico cuando se produce la resonancia de plasmón, como por ejemplo la diferencia de reflectividad de la luz con el campo aplicado en un sentido y en otro (TMOKE, del inglés *Transversal Magneto-Optical Kerr Effect*): cuando se produce el reconocimiento, la señal TMOKE se modifica (véase Fig. 1). Se obtiene así una mejora en el límite de detección, es decir, se consigue detectar concentraciones más bajas de analito [7,8].



Otra posibilidad es emplear nanoestructuras de oro en lugar de una lámina delgada, y por tanto la detección está basada en detectar cambios en la resonancia de plasmón localizado. La posición espectral de esa resonancia depende de la composición, tamaño y forma de las nanopartículas, así como también del índice de refracción del medio dieléctrico alrededor de ellas. La dependencia con los tres primeros factores se emplea para seleccionar en qué región del espectro se desea trabajar, ya en el espectro visible, ya en el infrarrojo cercano (sintonización espectral), lo que resulta útil para maximizar la respuesta de biosensibilidad de las nanoestructuras. Por último, las interacciones biomoleculares en la superficie de las nanoestructuras provocan un cambio local de índice de refracción, que se detecta por el desplazamiento de la resonancia [9].

Otros biosensores emplean efectos de transporte eléctrico, como por ejemplo cambios en la impedancia de una multicapa magnética cuando se aplica un campo magnético, fenómeno denominado magneto-impedancia gigante (GMI, *Giant Magneto-Impedance*). Se emplean partículas magnéticas en disolución convenientemente funcionalizadas para captar el analito deseado (un biomarcador que indique la presencia de una enfermedad: un fragmento de ADN, un anticuerpo, una proteína, una enzima...), y cuando se depositan sobre

la multicapa, el campo magnético dipolar de las partículas se detecta por los cambios en la respuesta de GMI de aquella [10].

Otro tipo de biosensores basados en nanomateriales son los biosensores nanomecánicos [11]. En su configuración más habitual, se trata de unas micropalanca con forma de trampolín y espesor del orden de 1000 nm (véase Fig. 2), que están funcionalizadas con la especie complementaria al analito que se desea detectar. Cuando se produce el reconocimiento molecular, cambian tanto la masa como las propiedades mecánicas del sistema (elasticidad, rigidez). Suele trabajarse en modo dinámico, es decir, aprovechando las resonancias mecánicas de la micropalanca (por ejemplo, el primer modo flexural) y empleando un láser que rebota en la micropalanca y llega a un fotodetector, para detectar cómo se mueve aquella.

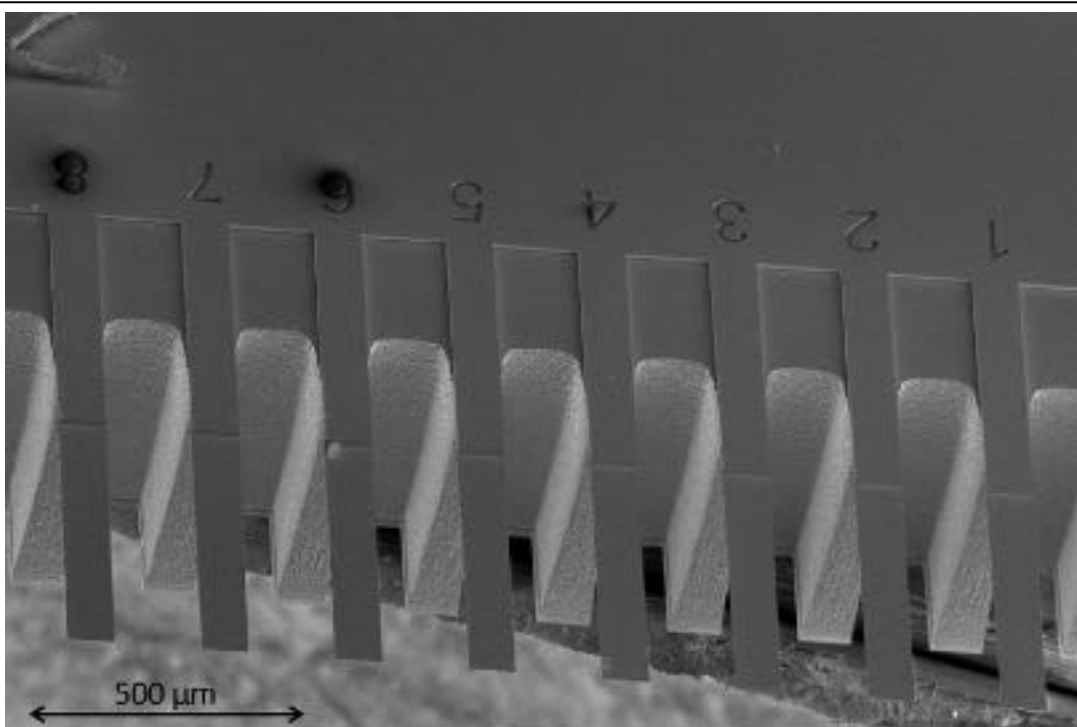


Fig.2. Imagen de SEM de un conjunto de micropalanca (cortesía del grupo de Bionanomecánica del Instituto de Micro y Nanotecnología, CSIC).

Mycobacterium tuberculosis, la bacteria responsable de la tuberculosis [12], y el VIH en un estadio muy incipiente de la enfermedad, apenas una semana después de la infección [13], han podido detectarse con estos biosensores nanomecánicos. De hecho, son tan sensibles que pueden ser una buena herramienta para establecer un diagnóstico a partir de las propiedades mecánicas de las células [14]. Por ejemplo, se sabe que las células de cánceres metastásicos son más elásticas que las células sanas, un hecho que está relacionado con la capacidad de aquéllas para invadir zonas distantes de nuestro organismo, puesto que tienen que atravesar canales muy estrechos, como por ejemplo finos capilares. Un conjunto de células cancerígenas sobre

una micropalanca daría por tanto una respuesta mecánica distinta a la que tendría un conjunto de células sanas.

Nanotecnología para tratamiento de enfermedades.

Aunque la nanotecnología puede emplearse para otros aspectos curativos importantes como son la regeneración de tejidos o la terapia génica, nos centraremos aquí en las soluciones que ofrece para paliar dos de los problemas de salud más acuciantes hoy en día: el cáncer, esto es, el crecimiento incontrolable de células con un comportamiento anormal, y las infecciones causadas por bacterias, dado que muchas de ellas desarrollan mecanismos adaptativos para resistir a los antibióticos convencionales.

Nanotecnología contra el cáncer.

El cáncer es actualmente la segunda causa de mortalidad en los países desarrollados, y se estima que una de cada tres personas deberá enfrentarse a esta enfermedad a lo largo de su vida. En la lucha contra el cáncer, hay planteadas diversas estrategias basadas en nanomateriales. Una de las más abordadas entre la comunidad científica es la hipertermia, que consiste en poner junto a las células cancerígenas unas nanopartículas que les aportan calor para destruirlas, ya sea mediante el desencadenamiento de la apoptosis o muerte celular programada cuando la temperatura alcanza unos 45 °C, ya sea mediante su carbonización o termoablación para temperaturas superiores a 50 °C. Existen dos vías para que las nanopartículas aporten calor, que se diferencian en el mecanismo y en el agente externo que se aplica:

1) La hipertermia plasmónica [15]. En este caso se emplean nanopartículas de metal noble que presentan una resonancia de plasmón localizada: cuando se las ilumina con luz de la longitud de onda a la que presentan la resonancia, que está en el rango espectral del infrarrojo, esa radiación es absorbida totalmente y la nanopartícula se calienta. Son muy efectivas las nanopartículas que presentan puntas donde se producen efectos de localización e intensificación de la luz, como nanobarras o nanoestrellas[16].

2) La hipertermia magnética [3]. En esta terapia se emplean nanopartículas magnéticas, en especial las de óxidos de hierro debido a su buena biocompatibilidad, y de tamaño suficientemente pequeño como para presentar un comportamiento superparamagnético: esto significa que, debido a las fluctuaciones térmicas, no tienen imanación estable en ausencia de campo magnético aplicado, por lo que se mantiene la estabilidad coloidal y se impide que se aglomeren por atracción magnética. Cuando se aplica un campo magnético alterno externo, las nanopartículas transfieren calor a su entorno, ya sea porque sus momentos magnéticos o espines se van orientando siguiendo las variaciones de ese campo y disipando energía, ya sea porque cada nanopartícula rota como un todo y fricciona con el medio circundante (esos mecanismos se denominan relajación de Néel y relajación de Brown, respectivamente).

Otra estrategia basada en nanopartículas magnéticas se fundamenta en una actuación magneto-mecánica. En esta ocasión, se emplean estructuras que sí presentan un estado ferromagnético, aunque con una pequeña

imanación neta remanente, y se aplica un campo alterno para que sea toda la nanoestructura quien rote, generando así una fricción mecánica que provoca la ruptura de las células cancerígenas a las que están adheridas o que incluso las han internalizado [17]. Para que el proceso sea más eficiente, se prefieren geometrías distintas a la esférica, por ejemplo nanodiscos [18,19], véase Fig. 3.

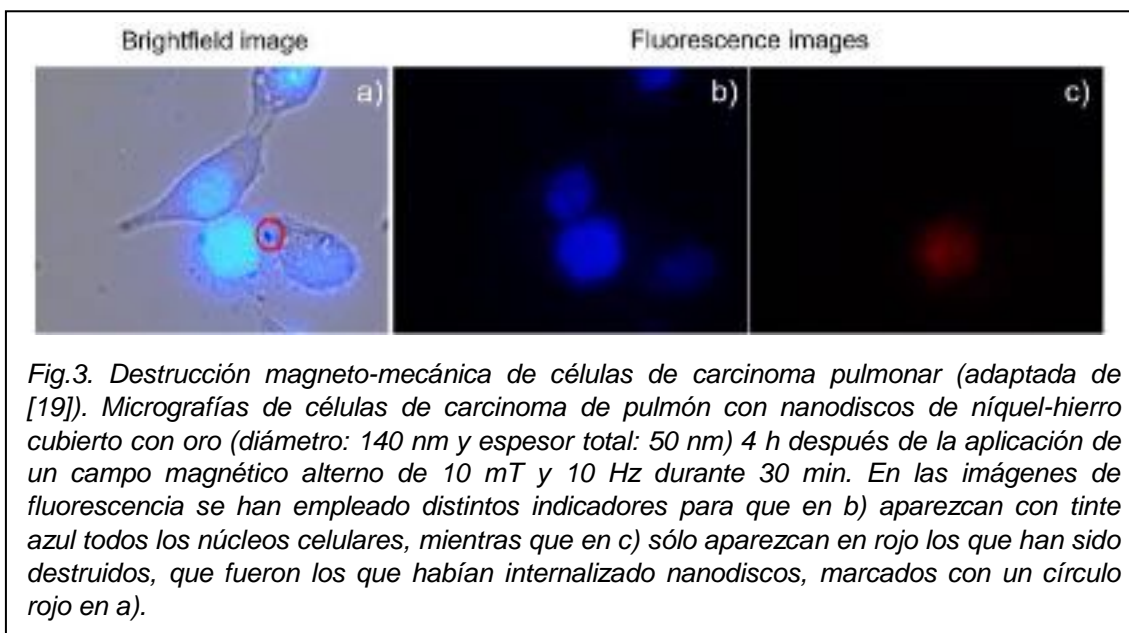


Fig.3. Destrucción magneto-mecánica de células de carcinoma pulmonar (adaptada de [19]). Micrografías de células de carcinoma de pulmón con nanodiscos de níquel-hierro cubierto con oro (diámetro: 140 nm y espesor total: 50 nm) 4 h después de la aplicación de un campo magnético alterno de 10 mT y 10 Hz durante 30 min. En las imágenes de fluorescencia se han empleado distintos indicadores para que en b) aparezcan con tinte azul todos los núcleos celulares, mientras que en c) sólo aparezcan en rojo los que han sido destruidos, que fueron los que habían internalizado nanodiscos, marcados con un círculo rojo en a).

También se investiga en el suministro dirigido de fármacos mediante nanopartículas: en este caso, las nanopartículas hacen de portadoras de medicamentos que administran localmente en la zona del organismo donde es necesario actuar, zona que podemos denominar “el blanco” al que se desea apuntar, que en muchos casos es un tumor cancerígeno o incluso orgánulos celulares específicos como el núcleo o la mitocondria. En esta ruta, las nanopartículas se funcionalizan para que lleven un fármaco adherido o incorporado, que se libera cuando haya alcanzado el blanco mediante la activación del mecanismo adecuado por un estímulo. Ese estímulo puede ser interno a la patología existente, como un cambio en el pH, una degradación enzimática o un potencial redox, o puede ser externo, como aplicar un campo magnético, ultrasonidos, o una determinada iluminación.

Es importante mencionar aquí otro concepto clave: la encapsulación, es decir, nanoentidades que ejercen de contenedores o cápsulas en cuyo interior se halla un fármaco. Entre las más estudiadas se encuentran las nanopartículas mesoporosas de sílice [20], los liposomas (vesículas esféricas de lípidos) [21] y sus análogos a base de polímeros [22], llamados polimerosomas. La principal ventaja que ofrecen estas dos últimas nanocápsulas es su mimetismo mejorado en medio fisiológico, por lo que pueden evitar el reconocimiento y la respuesta del sistema inmune.

Por último, debemos responder a una pregunta fundamental: ¿cómo se logra que, en todas estas estrategias (hipertermia, actuación magneto-mecánica, reparto local de fármacos) solo se vean afectadas las células cancerígenas y no las sanas? De nuevo, de forma más extendida, mediante

una funcionalización adecuada de las nanopartículas, por ejemplo acoplándoles anticuerpos que reconozcan los antígenos específicos de tumores (proteínas que aparecen solo en células tumorales). En el caso de nanopartículas magnéticas, si el tumor está localizado, pueden acumularse allí colocando un imán permanente en su proximidad [23]. Es necesario que ese imán genere un gradiente de campo suficientemente intenso como para vencer al movimiento del flujo sanguíneo, y hay que tener en cuenta además que el tejido existente entre el blanco y el imán también podría acumular nanopartículas: por ello, imanes externos solo podrán emplearse para tumores cercanos a la superficie corporal, mientras que tumores internos requerirán la implantación quirúrgica del imán (y su posterior extracción).

Asimismo, en determinados tumores se manifiesta el efecto de retención y permeabilidad aumentadas, también conocido como efecto EPR por sus siglas en inglés (*Enhanced Permeability and Retention*) [24]: sucede que el tejido tumoral presenta una vascularización inmadura con grandes fenestraciones o aberturas, lo que provoca que en estos tejidos se acumule de forma natural una cantidad de nanopartículas mucho mayor que en los tejidos sanos.

Nanotecnología contra las infecciones.

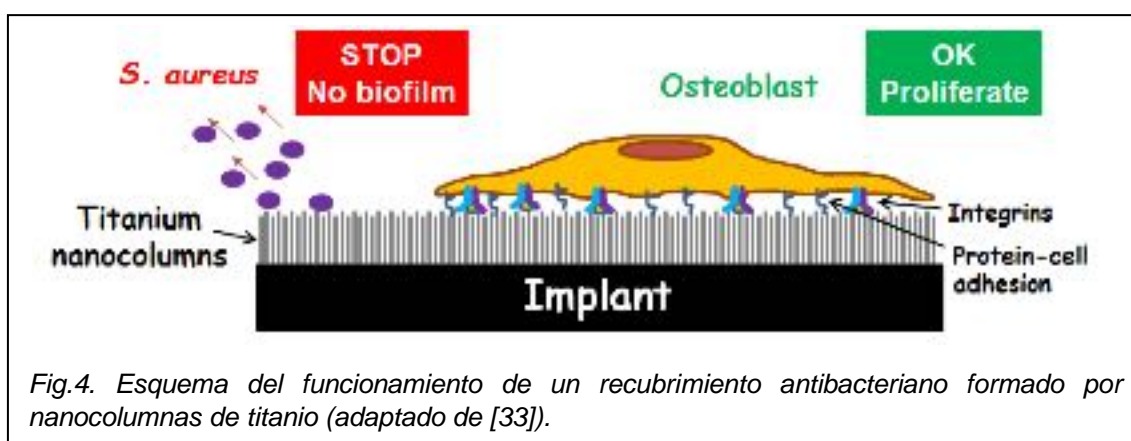
La aparición de bacterias resistentes a los fármacos es un fenómeno que causa gran preocupación hoy en día: por ejemplo, según el Centro de Control de las Enfermedades (CDC, *Center for Disease Control*) las infecciones bacterianas en Estados Unidos afectan a unos dos millones de personas anualmente, y de continuar la tendencia actual, se prevé que en 2050 morirán más personas debido a ellas que a todas las demás enfermedades juntas, incluido el cáncer. Por ello, hay una importante labor investigadora en este campo, en la que los nanomateriales tienen mucho que aportar.

En los últimos años se ha verificado la acción antibacteriana de diversas nanopartículas metálicas, en especial de plata [25], cobre [26] y zinc [27]. Estas nanopartículas se oxidan superficialmente de forma natural en condiciones ambientales o en medio fisiológico, y presentan efectos antibacterianos debido a dos mecanismos [28]. Por un lado, liberan iones metálicos que provocan daños funcionales en la membrana celular, alterando su integridad. Por otro, producen especies reactivas de oxígeno (ROS, por las siglas de *Reactive Oxygen Species*: iones oxígeno, radicales libres y peróxidos) que inducen estrés oxidativo y acaban provocando la apoptosis: este suele ser el mecanismo más importante [25]. Conviene señalar que estos efectos también se están explorando en la lucha contra el cáncer, y de hecho ya hay nanopartículas metálicas o de óxidos metálicos que presentan actividad antibacteriana y también anticancerígena [29].

Se trabaja también en el desarrollo de recubrimientos nanoestructurados y biocompatibles con propiedades antibacterianas, que paliarían los problemas de infección y rechazo en diversos biomateriales y herramientas médicas, tales como implantes dentales, implantes ortopédicos, marcapasos y catéteres. Existen tres estrategias posibles: 1) superficies que liberan fármacos antimicrobianos; 2) superficies que matan directamente las bacterias adheridas (eliminación por contacto); y 3) superficies que impiden la adhesión de las

bacterias. Debido a la gran capacidad adaptativa de las bacterias frente a los antibióticos, se consideran más interesantes a largo plazo estas dos últimas estrategias [30].

La eliminación por contacto puede lograrse incorporando nanoentidades con propiedades bactericidas, como por ejemplo nanopartículas de plata [31]. Por su parte, impedir la adhesión de las bacterias se obtiene con una corrugación en la nanoescala que hace que las bacterias tengan pocos puntos de anclaje accesibles [32]. En el caso de los recubrimientos para implantes, es un requisito imprescindible que esa corrugación no impida la proliferación de las células eucarióticas necesarias, como son los osteoblastos (las células que generan hueso) en los implantes ortopédicos. Afortunadamente, los tamaños son distintos: las bacterias tienen tamaños entre 500 nm y 2 micras, mientras que las células eucarióticas involucradas tienen tamaños de decenas de micras, por lo que con los parámetros adecuados pueden conseguirse superficies nanoestructuradas que inhiban la proliferación bacteriana sin afectar (o incluso mejorar) el desarrollo de las segundas. Por ejemplo, un recubrimiento a modo de tapiz formado por nanocolumnas de titanio de 50 nm de diámetro, 300 nm de largo y separaciones entre columnas de 100 a 200 nm disminuye notablemente la adhesión bacteriana e impide la formación del biofilm (impidiendo por tanto la infección), mientras que permite el normal desarrollo de los osteoblastos [33], véase Fig. 4.



Además, es posible mejorar las prestaciones de este recubrimiento sintetizando unas nanopartículas de teluro sobre ese tapiz [34]: se logra así una sinergia entre la baja adhesión bacteriana debida a las nanocolumnas de titanio y el efecto bactericida (eliminación por contacto debida a la producción de ROS) causado por esas nanopartículas de teluro, sin comprometer por ello la viabilidad de los osteoblastos, véase Fig. 5.

Consideraciones finales.

Acabamos de ver que los nanomateriales no solo deben ser efectivos para el diagnóstico o el tratamiento de una enfermedad, sino que también deben ser inocuos para las células sanas y que permitan su normal desarrollo. Los

estudios de biocompatibilidad deben contemplar además otros aspectos: es necesario comprobar que no generen rechazo inmunológico ni provoquen carcinogénesis, y considerar su posible transformación a lo largo de su ciclo de vida, analizando su toxicidad: a qué órganos pueden afectar, qué alteraciones pueden producir y a partir de qué umbrales. Asimismo, es importante estudiar cómo se eliminan del organismo y qué sucede con ellos posteriormente, teniendo en consideración su posible reciclaje.

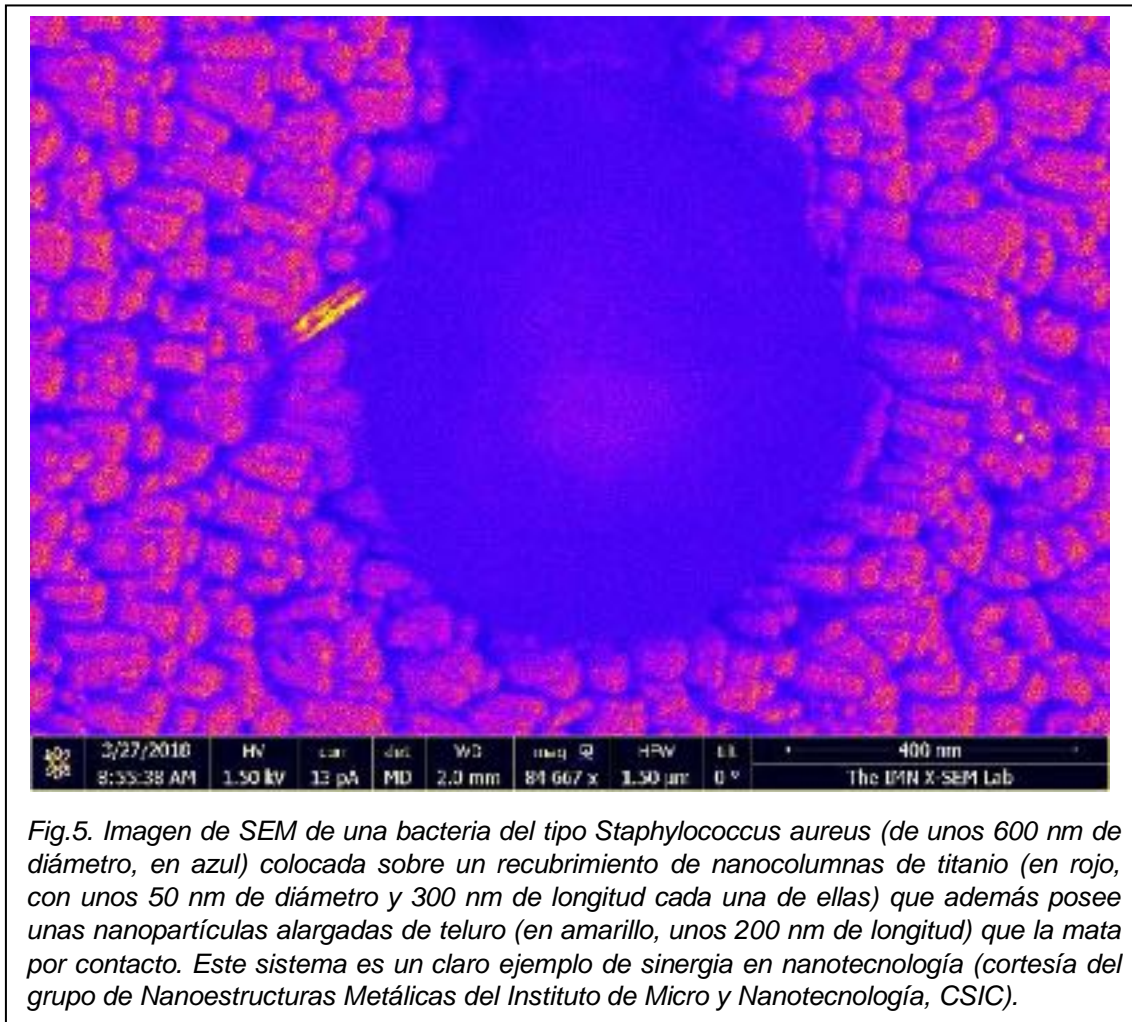


Fig.5. Imagen de SEM de una bacteria del tipo Staphylococcus aureus (de unos 600 nm de diámetro, en azul) colocada sobre un recubrimiento de nanocolumnas de titanio (en rojo, con unos 50 nm de diámetro y 300 nm de longitud cada una de ellas) que además posee unas nanopartículas alargadas de telurio (en amarillo, unos 200 nm de longitud) que la mata por contacto. Este sistema es un claro ejemplo de sinergia en nanotecnología (cortesía del grupo de Nanoestructuras Metálicas del Instituto de Micro y Nanotecnología, CSIC).

Debido a todo ello, hasta que se comercializa una terapia hay que recorrer un largo camino que dura varios años: se inicia con los ensayos *in vitro*, si son positivos se pasa normalmente a ensayos con animales y por último se realizan las distintas fases de los estudios clínicos. Por eso, le ruego al lector o a la lectora de este capítulo un poco de paciencia si no encuentra aún en el hospital algunas de las técnicas que aquí se han mencionado: le aseguro que no son ciencia ficción, y muchas de ellas acabarán estando disponibles para su uso.

Referencias:

- [1] Gas Phase Synthesis of Nanoparticles, Editor: Y. Huttel, Wiley-VCH Verlag GmbH, April 2017. ISBN 978-3-527-34060-6.
- [2] Green Synthesis of Metallic Nanoparticles via Biological Entities, Monaliben Shah, Derek Fawcett, Shashi Sharma, Suraj Kumar Tripathy and G errard Eddy Jai Poinern, *Materials* 8, 7278 (2015).
- [3] Biological applications of magnetic nanoparticles, M. Colombo, S. Carregal-Romero, M. F. Casula, L. Guti errez, M. P. Morales, I.B. B ohm, J. T. Heverhagen, D. Prospero, W. J. Parak, *Chemical Society Reviews*, 41, 4306 (2012).
- [4] Nanoprobes with near-infrared persistent luminescence for in vivo imaging, Q.M. de Chermont, C. Chaneac, J. Seguin, F. Pelle, S. Ma trejean, J.P. Jolivet, et al., *P. Natl. Acad. Sci. U. S. A.* 104, 9266 (2007).
- [5] A new near-infrared persistent luminescence nanoparticle as a multifunctional nanoplatform for multimodal imaging and c ancer therapy, Junpeng Shi, Xia Sun, Shenghui Zheng, Jinlei Li, Xiaoyan Fu, Hongwu Zhang, *Biomaterials* 152, 15 (2018).
- [6] Surface Plasmon Resonance Sensors for Detection of Chemical and Biological Species, J. Homola, *J. Chem. Rev.* 108, 462 (2008).
- [7] Enhanced antibody recognition with a magneto-optic surface plasmon resonance (MO-SPR) sensor, Maria Grazia Manera, Elias Ferreiro-Vila, Jose Miguel Garcia-Martin, Antonio Garcia-Martin, Roberto Rella, *Biosensors and Bioelectronics* 58, 114 (2014).
- [8] Propiedades magneto- pticas y plasm nicas entrelazadas en multicapas magnetoplasm nicas de metal y metal/diel ctrico. Tesis doctoral de El as Ferreiro Vila, Universidad de Santiago de Compostela (2012).
- [9] LSPR-based nanobiosensors, Borja Sep lveda, Paula C. Angelom , Laura M. Lechuga, Luis M. Liz-Marz n. *Nano Today* 4, 244 (2009).
- [10] An integrated giant magnetoimpedance biosensor for detection of biomarker, T. Wang, Z. Yang, C. Lei, J. Lei, Y. Zhou. *Biosensors and Bioelectronics* 58, 338 (2014).
- [11] Comparative advantages of mechanical biosensors, J.L. Arlett, E.B. Myers, and M.L. Roukes, *Nature Nanotechnology* 6, 203 (2011).
- [12] Label-Free DNA-Based Detection of Mycobacterium tuberculosis and Rifampicin Resistance through Hydration Induced Stress in Microcantilevers, Carmen M. Dom nguez, Priscila M. Kosaka, Alma Sotillo, Jes s Mingorance, Javier Tamayo, and Montserrat Calleja. *Anal.Chem.* 87, 1494 (2015).
- [13] Ultrasensitive detection of HIV-1 p24 antigen by a hybrid nanomechanical-optoplasmonic platform with potential for detecting HIV-1 at first week after infection, Priscila M. Kosaka, Valerio Pini, Montserrat Calleja, Javier Tamayo, *PLOS ONE*, 12: e0171899 (2017).

- [14] Mass and stiffness spectrometry of nanoparticles and whole intact bacteria by multimode nanomechanical resonators, O. Malvar, J. J. Ruz, P. M. Kosaka, C. M. Domínguez, E. Gil-Santos, M. Calleja, J. Tamayo, *Nature Communications*, 7, 13452 (2016).
- [15] Photo-thermal tumor ablation in mice using nearinfrared-absorbing nanoparticles, D. Patrick O'Neal, Leon R. Hirsch, Naomi J. Halas, J. Donald Payne, Jennifer L. West, *Cancer Letters* 209, 171 (2004).
- [16] Tailoring Surface Plasmons through the Morphology and Assembly of Metal Nanoparticles, Luis M. Liz-Marzan, *Langmuir* 22, 32 (2006).
- [17] Biofunctionalized magnetic-vortex microdiscs for targeted cancer-cell destruction, D. H. Kim, E. A. Rozhkova, I. V. Ulasov, S. D. Bader, T. Rajh, M. S. Lesniak and V. Novosad, *Nat Mater.* 9, 165 (2010).
- [18] Properties of Permalloy nanodiscs in magnetic vortex state and magneto-mechanical treatment of cancer cells, M. Goiriena-Goikoetxea, I. Orue, K. Guslienko, E. Berganza, M. Jaafar, A. Asenjo, D. Munoz, A. Muela, A. Garcia-Arribas, *Proceedings 2018 IEEE INTERMAG*, DOI: 10.1109/INTMAG.2018.8508850
- [19] Magnetic vortex nanodiscs for cancer cell destruction. Tesis doctoral de Maite Goiriena-Goikoetxea, Universidad del País Vasco (2017).
- [20] Mesoporous Silica Nanoparticles for Drug Delivery: Current Insights, M. Vallet-Regí, M. Colilla, I. Izquierdo-Barba and M. Manzano, *Molecules* 23, 47 (2018).
- [21] Biodistribution and pharmacokinetics of ¹⁸⁸Re-liposomes and their comparative therapeutic efficacy with 5-fluorouracil in C26 colonic peritoneal carcinomatosis mice, Tsai et al., *Int. J. Nanomedicine* 6, 2607 (2011).
- [22] Polymeric nanoparticles for targeted treatment in oncology: current insights, R. H. Prabhu, V. B. Patravale, M. D. Joshi, *Int. J. Nanomedicine* 10, 1001 (2015).
- [23] Simultaneous hyperthermia-chemotherapy with controlled drug delivery using single-drug nanoparticles, I. Sato et al., *Sci. Rep.* 6, 24629 (2016).
- [24] Macromolecular therapeutics in cancer treatment: The EPR effect and beyond, H. Maeda, *Journal of Controlled Release* 164, 138 (2012).
- [25] Role of reactive oxygen species in the antibacterial mechanism of silver nanoparticles on *Escherichia coli* O157:H7, Xu et al., *Biometals* 25, 45 (2012).
- [26] Antimicrobial Properties of Copper Nanoparticles and Amino Acid Chelated Copper Nanoparticles Produced by Using a Soya Extract, DeAlba-Montero et al., *Bioinorg. Chem. Appl.* 2017, 1064918 (2017).
- [27] Antibacterial activity of silver and zinc nanoparticles against *Vibrio cholerae* and enterotoxigenic *Escherichia coli*, Salem et al., *Int. J. Med. Microbiol.* 305, 85 (2015).
- [28] The antimicrobial activity of nanoparticles: present situation and prospects for the future, L. Wang, C. Hu, L. Shao, *Int. J. Nanomedicine* 12, 1227 (2017).
- [29] Synthesis and characterization of PVP-coated telluriumnanorods and their antibacterial and anticancer properties. Christopher D. Brown, David M. Cruz, Amit K. Roy, Thomas J. Webster, *J. Nanopart. Res.* 20: 254 (2018).
- [30] In vitro methods for the evaluation of antimicrobial surface designs, Jelmer Sjollema, Sebastian A.J. Zaat, Veronique Fontaine, Madeleine Ramstedt, Reto Luginbuehl, Karin Thevissen, Jiuyi Li, Henny C. van der Mei, Henk J. Busscher, *Acta Biomaterialia* 70, 12 (2018).
- [31] Antibacterial effects and biocompatibility of titanium surfaces with graded silver incorporation in titania nanotubes, Mei et al., *Biomaterials* 35, 4255 (2014).
- [32] Antibacterial titanium nanopatterned arrays inspired by dragonfly wings, Chris M. Bhadra, Vi Khanh Truong, Vy T. H. Pham, Mohammad Al Kobaisi, Gediminas Seniutinas, James Y. Wang, Saulius Juodkazis, Russell J. Crawford and Elena P. Ivanova, *Sci.Rep.* 5, 16817 (2015).
- [33] Nanocolumnar coatings with selective behavior towards osteoblast and *Staphylococcus aureus* proliferation, Isabel Izquierdo-Barba, José Miguel García-Martín, Rafael Álvarez, Alberto Palmero, Jaime Esteban, Concepción Pérez-Jorge, Daniel Arcos, María Vallet-Regí. *Acta Biomaterialia* 15, 20 (2015).
- [34] Synergic antibacterial coatings combining titanium nanocolumns and tellurium nanorods, David Medina-Cruz, María Ujué González, William Tien-Street, Marcial Fernández-Castro, Ada Vernet-Crua, Iván Fernández-Martínez, Lidia Martínez, Yves Huttel, Thomas J. Webster and José Miguel García-Martín. *Nanomed.-Nanotechnol. Biol. Med.* 17, 36 (2019).

José Miguel García-Martín

Doctor en Ciencias Físicas.

Investigador Científico, Instituto de Micro y Nanotecnología, CSIC.

Capítulo 24

Sentir el pulso del Universo: ¿qué podemos esperar aprender de las ondas gravitatorias?

Ver con luz o ver con otras 'luces'.

No es ni aventurado ni exagerado decir que durante la práctica totalidad de la historia humana, el conocimiento del Universo, más allá de nuestro alcance inmediato, nos ha llegado mediante la luz. Primero a ojo desnudo, y luego ayudados de instrumentos: anteojo de Galileo, telescopios, registro en placas fotográficas,... No teníamos ninguna razón para suponer que el Universo estuviera emitiendo 'otras luces'. Ahora sabemos que lo hace. Y además por partida doble.

A mediados de los 1860s el físico escocés J.C. Maxwell predijo la existencia de ondas electromagnéticas y adelantó la hipótesis de que la luz, de naturaleza entonces desconocida, era una onda electromagnética. Comprobada hacia 1890 su existencia por Hertz, resulta haber tales ondas en todo un amplio espectro de frecuencias, desde las de radio hasta los rayos X y los rayos gamma; la luz visible ocupa solo una estrecha región intermedia de ese espectro.

Las aplicaciones prácticas de semejante descubrimiento, inimaginables antes, llegaron muy pronto: en los 1920s comenzaron las emisiones comerciales de radio, y poco después de televisión, utilizando las regiones del espectro en las que es relativamente fácil producir estas ondas.

Ya que desde el universo nos llega luz, no tardó en surgir la pregunta de si acaso no nos llegarían también 'luces' en otras bandas de frecuencia. Cuando se descubrió que las ondas cortas de radio emitidas en la superficie de la Tierra se reflejaban en una capa alta de la atmósfera, se pensó que esa capa también reflejaría hacia afuera las eventuales ondas de radio procedentes del espacio, lo que impediría su eventual observación. Ahora sabemos que la atmósfera no bloquea todas las zonas del espectro de radio: la primera observación casual, en los 1930s, de ondas de radio procedentes del cosmos promueve el desarrollo de la radioastronomía, que alcanza en los 1950s su mayoría de edad permitiendo en las décadas siguientes 'ver' el Universo con nuevos ojos, los de los radiotelescopios.

Una vez entendido que la luz visible no es el único mensajero del Universo, era natural e inevitable intentar observar el Universo en otras frecuencias: microondas, infrarrojo, ultravioleta, rayos X y rayos Gamma, lo que requirió entre 1950 y 1990 el diseño de instrumentación específica, que cuando estuvo disponible mostró que recibimos señales del Universo en todas esas

bandas. Y a partir de los 1990s hemos podido situar algunos de esos instrumentos en el espacio, desde donde muchas de las dificultades ligadas con la observación terrestre desaparecen. Ahora estamos 'viendo', desde todas esas 'ventanas', una imagen incomparablemente más completa del Universo.

Las ondas gravitatorias: sentir el pulso del Universo.

En los últimos 30 años hemos ido acumulando evidencias indirectas de la existencia de *otras ondas*, las ondas gravitatorias. Pero tuvimos que esperar a 2015, casi un siglo después de haber sido previstas, para observar directamente por vez primera la radiación gravitatoria. Con esa detección, en vez de 'ver' el Universo, podemos 'sentir' su pulso oscilante. La astronomía de ondas gravitatorias, ahora en su infancia, es más que otra nueva ventana al Universo, ya que nos permitirá acceder a información vedada por medios ópticos.

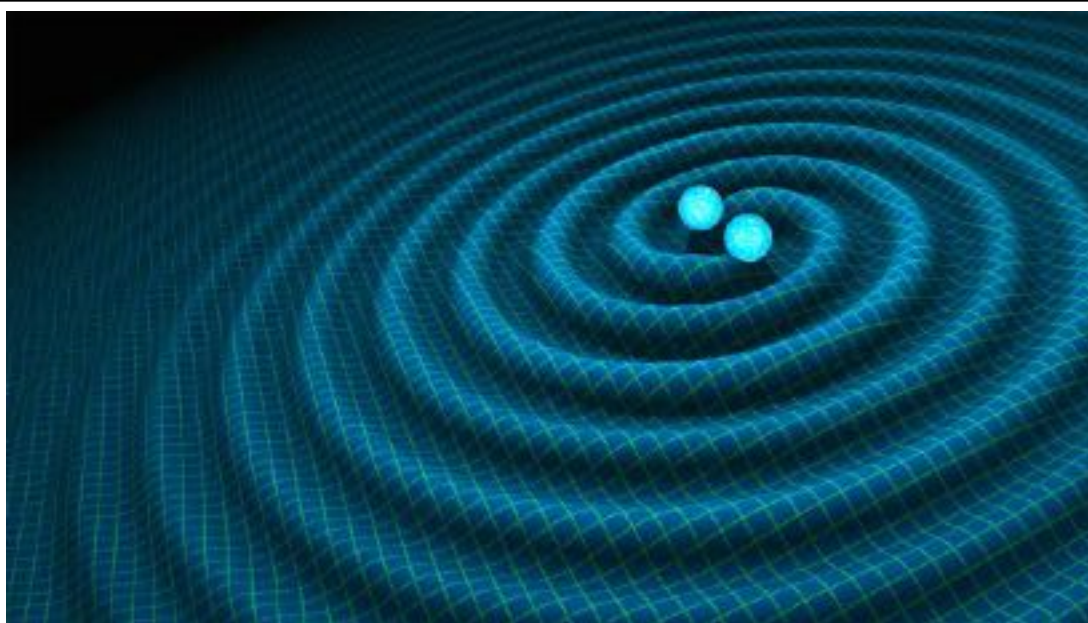


Fig. 1. Una imagen artística de las ondas gravitatorias generadas por un sistema binario de estrellas de neutrones. Crédito: R. Hurt/Caltech-JPL.

<https://www.jpl.nasa.gov/news/news.php?feature=5137>

Previstas inicialmente por Einstein en 1916 como una consecuencia de su teoría de la gravedad, luego se llegó a dudar de que esas ondas fueran más que una solución matemática sin un correlato físico real. Entre 1950 y 1960 se concluye que esas dudas no están justificadas, y que tales ondas deben existir. Lo que se ignoraba por completo en 1960 era si en el Universo habría fuentes capaces de emitir ondas gravitatorias de intensidad suficiente como para poder ser detectadas. Pues bien pudiera ocurrir que, existiendo, esas ondas fueran tan extremadamente débiles que su detección fuera imposible en la práctica.

¿Qué es una onda gravitatoria?: analogías y diferencias con las ondas electromagnéticas.

La radiación electromagnética se produce en las transiciones atómicas y también durante el movimiento acelerado de cargas eléctricas (por ejemplo átomos ionizados), y puede automantenerse en regiones sin cargas ni corrientes eléctricas por un mecanismo de 'realimentación' perfectamente descrito en las ecuaciones de Maxwell: un campo eléctrico (o magnético) variable con el tiempo genera un campo magnético (o eléctrico); esas variaciones son las ondas que viajan a cierta velocidad universal.

En contraste con las ondas electromagnéticas, que se producen en átomos o en moléculas individuales, la radiación gravitatoria se origina, de forma coherente, por movimientos acelerados de enormes cantidades de materia. La onda, una vez producida, se propaga en el espacio a la misma velocidad universal, la de la luz. Una onda gravitatoria también está caracterizada por su frecuencia y por su amplitud, que disminuye con la distancia a la fuente de la onda con el mismo factor $1/r$ con el que decrece la amplitud de las ondas electromagnéticas. Pero hay dos diferencias esenciales.

La primera es que no hay rayas espectrales gravitatorias, que en el caso electromagnético son la herencia de la producción de la radiación por átomos o moléculas individuales. En contraste, la frecuencia de la radiación gravitatoria está determinada por la frecuencia del movimiento, típicamente de rotación, de las grandes cantidades de materia / energía que producen la onda. De fuentes de diferentes tipos debemos esperar recibir ondas gravitatorias de diversas frecuencias, que además nos llegarán afectadas, al igual que las electromagnéticas, del desplazamiento 'al rojo' debido a la expansión del Universo.

La segunda gran diferencia es que la materia ordinaria es transparente a la radiación gravitatoria: una onda gravitatoria que llegue a la Tierra la atravesará limpiamente y seguirá su propagación. Lo que al paso de estas ondas oscila es *el propio espacio-tiempo*, que sufre una contracción y expansión periódica y alternada en las dos direcciones perpendiculares a la dirección de propagación. La amplitud de esa onda es el alargamiento o acortamiento relativo que se produce en el propio espacio: incluso para ondas gravitatorias bastante intensas, la amplitud de este 'latido' del espacio es casi imperceptible, por ejemplo de una milésima de milímetro en un kilómetro, lo que se suele expresar diciendo que el espacio es extremadamente rígido.

En los 1950s, la única fuente conocida de ondas gravitatorias de cierta amplitud eran los sistemas binarios de estrellas. Las dos estrellas, mientras orbitan, emiten de manera permanente ondas gravitatorias con frecuencia el doble de la frecuencia orbital del sistema y amplitud extraordinariamente pequeña. Tales ondas transportan energía, que pierde el sistema binario, con lo que la separación entre las dos estrellas disminuye. En las últimas revoluciones del sistema binario, cuando las estrellas se acercan irremisiblemente girando con una frecuencia de rotación creciente (por conservación del momento angular), la amplitud de las ondas gravitatorias emitidas aumenta siguiendo un patrón característico. Al llegar a nosotros, esa

onda provoca una oscilación en el propio espacio de extrema pequeñez: para una onda emitida en esa fase final de colapso de dos estrellas como el Sol, supuesto ocurrido en nuestra vecindad en la galaxia, al paso de tal onda gravitatoria por la Tierra 'sentiríamos' que en una separación de 1km la amplitud de la oscilación de esa separación sería muy inferior al tamaño de un núcleo atómico.

Añadamos a eso que la tasa temporal de ocurrencia de tales fusiones dentro de nuestra galaxia es probablemente de solo una cada varios siglos o milenios, y que la parte más intensa de la onda gravitatoria, la emitida en las últimas revoluciones del sistema es pasajera y duraría solamente unas horas. Se entiende que durante bastante tiempo fuera dominante la opinión de que las ondas gravitatorias permanecerían siendo indetectables para siempre.

El cambio observacional.

En torno a 1960 algunas cosas comenzaron a cambiar en esa opinión con el descubrimiento, mediante radioastronomía, de nuevos objetos en el Universo.

En la década de 1950 se descubrieron unas fuentes con una emisión muy intensa en radio pero cuya naturaleza era misteriosa; se bautizaron 'quasars' ('quasi-estelares').

En 1967 se descubrieron otros objetos, los pulsars, que emitían señales de radio periódicas, con periodos del orden de segundos. A diferencia de los quasars, cuya naturaleza tardó en entenderse, la de los pulsars se estableció pronto: son sistemas binarios en los que uno de los miembros es una estrella de neutrones que emite un haz electromagnético colimado y muy intenso que gira, cual faro cósmico, con la rotación de la estrella.

En esas mismas fechas se asentó definitivamente el convencimiento, por argumentos indirectos, de la existencia real de los agujeros negros, sobre lo que hasta entonces no había consenso. Hoy sabemos que los quasars son emisiones electromagnéticas producidas por la materia que rota y va cayendo en los discos de acreción alrededor los agujeros negros supermasivos que hay en los núcleos galácticos.

¿Por qué fueron estos descubrimientos cruciales en nuestra historia? Porque un sistema binario de estrellas de neutrones o de agujeros negros es fuente de radiación gravitatoria muchísimo más intensa que la producida por un sistema binario de estrellas ordinarias. Esa radiación alcanza su máxima amplitud, bastante mayor que la de un sistema binario de estrellas de la misma masa, en el momento final de la fusión de los dos integrantes. Incluso si esa fusión ocurre en una galaxia relativamente alejada, la amplitud de la onda generada causaría, al llegar a la Tierra, una oscilación del orden de una milésima parte del tamaño del núcleo atómico en una separación del orden del kilómetro.

Cuando lo insignificante resulta ser significativo.

Esta amplitud de oscilación es comparable a la que causaría la fusión de un sistema binario estelar en nuestra vecindad galáctica. Pero la tasa de ocurrencia de procesos de fusión de agujeros negros o estrellas de neutrones, ahora en un volumen que incluye muchísimas galaxias, se estima en varios sucesos al año. Y además las frecuencias típicas de las ondas emitidas en esas fusiones son también mucho mayores, alrededor de los 100 Herzios. Lo que hace que el panorama de una posible detección cambie sustancialmente.

Si consiguiéramos construir un instrumento capaz de detectar directamente, en ese rango de frecuencias, una variación tan minúscula en el espacio entre dos puntos 'fijos', ese instrumento, funcionando a lo largo de un tiempo de observación razonable (digamos un año) debiera captar unos pocos sucesos. Esta fue la filosofía inicial que condujo a LIGO.

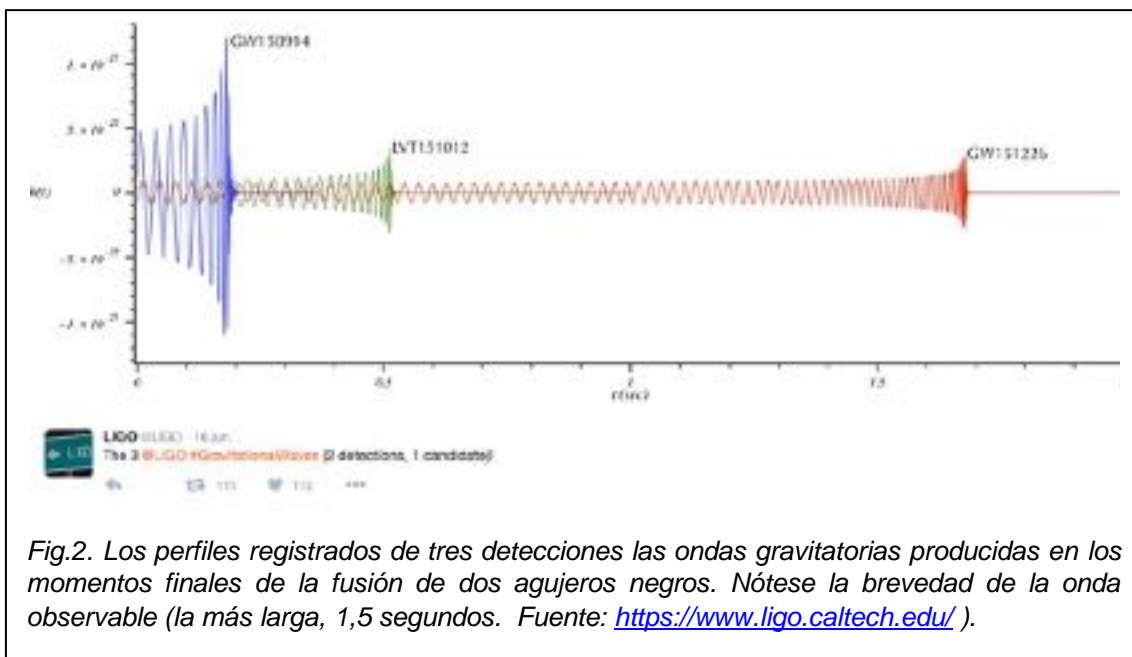


Fig.2. Los perfiles registrados de tres detecciones las ondas gravitatorias producidas en los momentos finales de la fusión de dos agujeros negros. Nótese la brevedad de la onda observable (la más larga, 1,5 segundos. Fuente: <https://www.ligo.caltech.edu/>).

El 'sentido común' nos diría que detectar una variación tan insignificante en la separación es una misión imposible. La historia subsiguiente es la de una auténtica epopeya, iniciada por unos pocos pioneros, que convencidos de que la detección sería posible, se atrevieron a diseñar, conseguir fondos, construir y optimizar un instrumento capaz de detectar tales minúsculas variaciones en la separación entre los espejos de un interferómetro. Esta es la historia de LIGO.

Detectores terrestres: LIGO.

El Observatorio Gravitacional por Interferometría Laser, LIGO, construido en la última década del S. XX y en funcionamiento desde hace quince años, es un instrumento para la observación de ondas gravitatorias mediante interferometría óptica. Esa técnica extremadamente sensible es capaz de detectar variaciones de fase relativas muy pequeñas entre dos haces de luz

que han tenido un origen común, se han separado siguiendo caminos diferentes y finalmente se reúnen de nuevo. Desde el interferómetro original de Michelson, empleado hace más de un siglo para tratar de detectar el posible movimiento de la Tierra a través del éter, las mejoras en la sensibilidad de estos instrumentos han sido espectaculares.

Hay varios efectos (ruido sísmico, térmico y fotónico) imposibles de sortear, y que, para un detector interferométrico en la superficie terrestre limitan (actualmente) las observaciones de ondas gravitatorias al rango de frecuencias entre 30Hz y 400Hz. Que afortunadamente son las de las ondas gravitatorias producidas en la fusión de dos agujeros negros de masa estelar o de dos estrellas de neutrones. En ese rango LIGO es capaz de detectar diferencias de las distancias recorridas por los dos haces del interferómetro del orden de la milésima parte del 'tamaño' de un núcleo atómico.

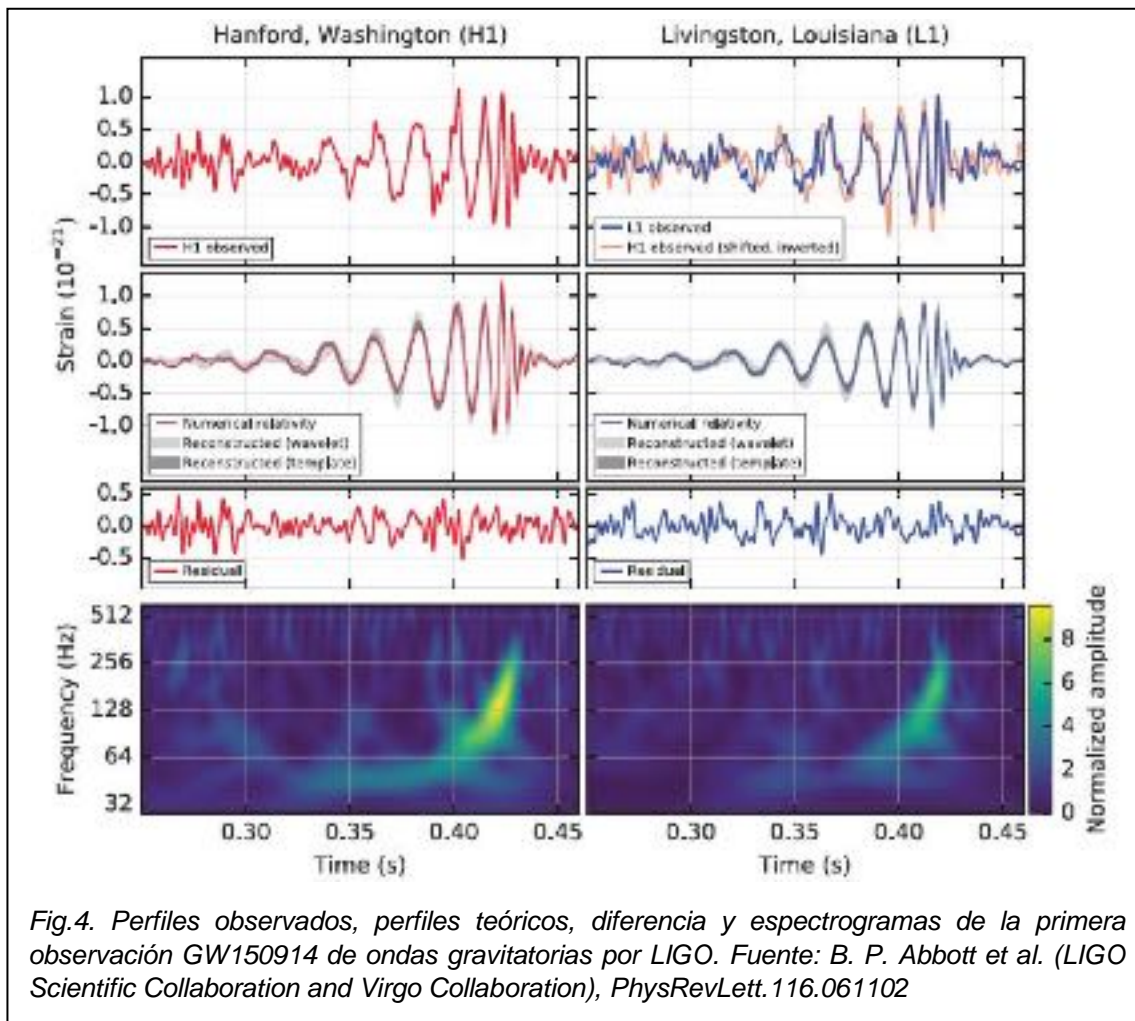


*Fig.3. Vista aérea de uno de los dos observatorios LIGO, en Hanford, EEUU.
Fuente: <https://www.ligo.caltech.edu/>*

LIGO consta de dos interferómetros gemelos, construidos en dos lugares de EEUU separados por 3000 km, y que funcionan en tándem, lo que permite eliminar señales espúreas que pudieran afectar solo a uno de ellos (olas del mar rompiendo en una costa alejada, zumbido sísmico originado por el centrifugado de una lavadora cercana). Esta duplicidad permite separar esas señales espúreas de las ondas gravitatorias procedentes del espacio, que afectarán a los dos interferómetros por igual.

En paralelo temporal con todo el desarrollo de LIGO, hay funcionando otros detectores de tipo interferométrico: GEO600 en Alemania, y VIRGO en Italia.

LIGO tuvo una larga fase de rodaje, en la que deliberadamente no se pretendía observar ninguna señal sino ir optimizando un instrumento de esa complejidad, en el que extraer las señales del ruido de fondo requiere además técnicas matemáticas muy elaboradas: inferencia bayesiana, filtrado de datos con plantillas precalculadas, Cuando en Septiembre de 2015 se inició la primera tanda de observaciones 'reales' se esperaba haber alcanzado la sensibilidad necesaria para detectar ondas gravitatorias, y en efecto, en los primeros días de observación se detectó GW150914, una señal 'de libro', la onda gravitatoria producida en las etapas finales de fusión de un sistema binario de agujeros negros.



Hasta Diciembre de 2018, y en poco más de un año de tiempo total de observación, hay ya confirmadas por parte de LIGO y de la colaboración LIGO-Virgo un total de diez observaciones de ondas gravitatorias generadas en la fusión de dos agujeros negros de masa estelar, y una de la fusión de dos estrellas de neutrones. La detección directa de las ondas gravitatorias generadas por un sistema binario de dos agujeros negros, aparte de su valor como tal, tiene otro adicional: es la comprobación más directa posible de la existencia de agujeros negros.

LIGO también podría llegar a detectar otras ondas gravitatorias de amplitud menor pero que en vez de ser ocasionales, como las anteriores, se produzcan de manera continuada, como las emitidas por un sistema binario durante su 'baile' antes del colapso; esto requiere registrar las observaciones a lo largo de bastante tiempo y analizarlas con técnicas muy elaboradas, un programa en curso.

Especulación para el futuro.

Sabiendo cómo hemos llegado hasta aquí, podemos especular sobre lo que, posiblemente, nos deparará el futuro. Este tipo de ejercicio conlleva un indudable riesgo. Pero no se trata ahora de aventurar qué descubrimientos nuevos en física nos esperan a la vuelta de unas décadas, algo a lo que desde luego yo no me atrevería. Hablando solo de ondas gravitatorias y una vez confirmada por observación directa su existencia, se trata de hacer un cierto ejercicio de prospectiva realista, describiendo los proyectos en curso para mejorar su sensibilidad, y adelantando lo que podemos esperar obtener de las previsibles observaciones con la tecnología disponible en un futuro cercano.

El futuro: más observatorios interferométricos.

Disponer de más observatorios, trabajando conjuntamente, permitirá precisar mucho más tanto el perfil de las ondas gravitatorias observadas en esos procesos astrofísicos, como la ubicación en la esfera celeste de su fuente. Actualmente hay en curso de construcción dos observatorios, LIGO-India, y KAGRA en Japón. En la tanda de observaciones de LIGO-Virgo que se iniciará en la primavera de 2019 se espera mejorar la sensibilidad y quizás, que KAGRA pueda incorporarse a lo largo de esa tanda, en la que la tasa de detecciones se estima en más de una por mes. Cuando todos los observatorios entren en funcionamiento conjunto, la detección de ondas gravitatorias se transformará, esperamos, en una herramienta rutinaria para la astronomía que probablemente nos deparará sorpresas.

¿Se necesitan realmente más observatorios? O ¿observatorios de otros tipos?

¿Puede LIGO detectar las ondas gravitatorias generadas por la caída de materia, siguiendo los discos de acreción, hacia los agujeros negros supermasivos? O ¿las ondas gravitatorias generadas por una sola estrella de neutrones rotante? La respuesta es 'no', ya que, incluso a igualdad de amplitudes, las frecuencias de estas ondas gravitatorias están muy por debajo del límite inferior de frecuencias al que LIGO es sensible.

Y es que las ondas gravitatorias generadas en diferentes procesos astrofísicos pueden tener frecuencias muy diferentes, desde un kilohertzio en el extremo superior hasta un extremo inferior del orden del attoherzio (cuyo

periodo es la vida entera del universo). LIGO solo puede 'sentir' en la muy pequeña parte superior de ese espectro. Para explorar debidamente el Universo en todas las frecuencias del espectro gravitatorio se necesitan más instrumentos. En cada rango de frecuencias es necesario llegar a una sensibilidad que permita detectar las amplitudes típicas esperables para las ondas de esas frecuencias; como además la amplitud de las ondas decrece con la distancia como $1/r$, duplicar la sensibilidad de cualquier método incrementa en un factor $2^3 = 8$ el volumen de Universo que estamos explorando efectivamente.

No mencionaremos otro tipo de detectores terrestres, los de barra resonante, cuyo desarrollo comenzado en los 1960 ha continuado hasta hoy, aunque de momento sin éxito en la detección.

KAGRA y el Telescopio Einstein.

Una posibilidad de mejora de los detectores interferométricos terrestres es hacerlos subterráneos, para que la roca situada por encima apantalle muchos de los ruidos sísmicos naturales o de origen humano. A primera vista puede parecer chocante que una buena ubicación para un observatorio de ondas procedentes del espacio sean túneles subterráneos, pero esto se debe a que la materia es 'transparente' a las ondas gravitatorias. El observatorio KAGRA, que próximamente entrará en operación regular es ya subterráneo.

Un plan más ambicioso, planteado en Europa y actualmente en fase de proyecto, es el 'Telescopio Einstein', que a pesar de la connotación óptica del nombre realmente es un gigantesco detector de ondas gravitatorias interferométrico con brazos de 10 kilómetros. La situación subterránea y la configuración de triángulo equilátero permitirán superar la sensibilidad de LIGO y proporcionar una ubicación más precisa de la fuente en la esfera celeste. Se espera que el Telescopio Einstein permita la detección de radiación gravitatoria producida por ciertos pulsares, o por sistemas binarios de enanas blancas, que son sistemas en los que los campos gravitatorios son intermedios entre los de los sistemas binarios de estrellas de neutrones y los de estrellas ordinarias.

Detectores en el espacio: LISA.

¿Es realista imaginar detectores de tipo interferométrico con brazos mucho mayores de 10Km en la Tierra? Probablemente no. Pero hay mucho espacio ahí fuera. Poner en funcionamiento un detector de tipo interferométrico en el espacio es el proyecto LISA (Laser Interferometer Space Antenna). En su concepción inicial, planteada conjuntamente por la agencia Espacial Europea, ESA, y la NASA, se trataba de tres satélites en órbita alrededor del Sol, formando un triángulo con lados del orden de 5 millones de kilómetros, conectados por haces de laser. Los detalles de la parte 'astrodinámica' del diseño son realmente fascinantes: los tres satélites estarían en órbita alrededor del Sol, con tres órbitas cuidadosamente seleccionadas para mantener de manera automática una configuración triangular. Los detalles del aspecto

interferométrico no lo son menos, y son bastante diferentes de los de un interferómetro terrestre, con las ventajas y los inconvenientes que resultan de una separación entre satélites tan monstruosamente grande y la dificultad añadida de unos sistemas que requiriendo una extrema precisión cuando estén en órbita, deben sobrevivir indemnes a un lanzamiento desde la Tierra que es cualquier cosa menos suave.

Una vez en órbita, los satélites que transportarán el interferómetro de LISA estarán sometidos a pequeños efectos que les infunden ligeras desviaciones sobre la caída libre que siguen en sus órbitas. Pero las masas que juegan el papel de espejos del interferómetro configurando el triángulo y cuya variación de separación se mide con los láseres deben estar en caída libre con una precisión muchísimo mayor. De no corregirse, esos efectos sobre los satélites harían inviable la finísima observación interferométrica requerida. ¿Cómo superar tal reto? La idea es que las masas vayan 'flotando' en el interior del satélite, como en una imaginaria jaula, con sensores que detectan cuando la masa comienza a aproximarse a uno de los lados de la 'jaula' (lo que habrá ocurrido como consecuencia de cualquier pequeña fuerza adicional a la gravedad que ha actuado sobre el satélite) y en ese momento corrigen la trayectoria del satélite sin actuar de ninguna manera sobre la masa que permanece 'flotando' en el interior de la jaula. Se consigue así que la masa vuelva a estar en reposo relativamente a la jaula; en resumen, la masa permanece siempre realmente en una caída libre casi perfecta, totalmente aislada de las pequeñas fuerzas no gravitatorias de toda índole que inevitablemente actuarán sobre el satélite.

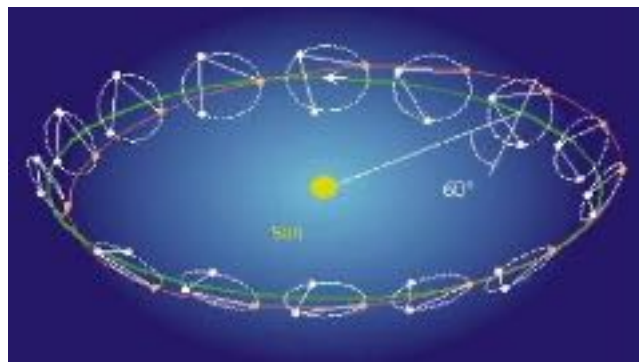


Fig.5. Esquema de las órbitas a lo largo de un año de los tres satélites que formarán LISA. La línea verde es la órbita de la Tierra, y el centro del triángulo sigue una trayectoria que casi coincide, aunque avanzada, con respecto a la de la Tierra.

Fuente: ESA, <http://sci.esa.int/lisa/35596-schematic-of-lisa-orbit/>

En 2011 la NASA se retiró del proyecto LISA, que afortunadamente fue mantenido por la ESA. Con la única finalidad de comprobar si la tecnología recién descrita para asegurar la caída libre de las masas era lo suficientemente precisa y fiable para un experimento a la escala en la que LISA debe funcionar, la ESA preparó la misión LISA Pathfinder. El éxito de ésta misión, ya completada, abrió el camino a la aprobación por parte de la ESA en 2017 de la misión completa, que en la década de 2030 debería estar en funcionamiento, con tres satélites separados 'solo' por 1,5 millones de kilómetros (cuatro veces la distancia a la luna).

Puede parecer que ese tipo de instrumento será 'mejor' que los terrestres; no es mejor, sino diferente. Y es que un observatorio así en el espacio también está limitado a una zona del espectro de frecuencias, con

limitaciones de naturaleza diferente a las que se encuentran en la Tierra. El rango efectivo de las frecuencias a las que LISA será sensible cubre tres órdenes de magnitud centrados en el miliherzio, que corresponde a ondas con periodos de un cuarto de hora. Ondas con esas frecuencias pueden ser producidas por la coalescencia final de sistemas binarios de estrellas ordinarias en nuestra galaxia, o por la captura de estrellas compactas o de agujeros negros estelares por los agujeros negros supermasivos que existen en todos los núcleos galácticos: como se ve, fuentes diferentes de las que puede detectar LIGO. Probablemente no veremos con LISA nada de lo que ya hayamos visto con LIGO, y nada de lo que veamos con LISA lo habremos visto con LIGO.

La puntualidad de los pulsars.

Las frecuencias de las ondas gravitatorias producidas por la coalescencia de agujeros negros supermasivos, como los que hay en el centro de las galaxias, están varios órdenes de magnitud por debajo del miliherzio, con períodos del orden de años (aquí de nuevo, la vieja y fundamental idea, contrabajo vs. violín, de que sistemas más grandes están asociados a frecuencias más bajas). Incluso LISA será incapaz de detectar en esas frecuencias, para las que se ha propuesto otro método que apenas requiere nueva tecnología, aunque sí un programa de observación de radioastronomía mantenido a largo plazo. Se trata del seguimiento temporal con gran precisión de los periodos de los llamados pulsars de milisegundo. ¿De qué va éste método?

Los pulsars, descubiertos en 1967, son estrellas de neutrones rotantes, con un intenso campo magnético, que emiten un haz de radiación electromagnética muy estrecho en una dirección que rota con la estrella. Si ese haz, en su movimiento de rotación, barre a la Tierra, lo que nosotros detectamos es un pulso brevísimo, que se repite en cada rotación del haz. En breve, un faro cósmico. Los periodos típicos con los que recibimos estos pulsos van desde unos pocos segundos hasta el milisegundo, y nos informan del periodo de rotación de la estrella de neutrones. Los pulsares de rotación más rápida son extraordinariamente estables, y pueden calificarse como los relojes más precisos del Universo, más que los mejores relojes atómicos actuales.

Imaginemos que una onda gravitatoria de período del orden de años atraviesa la zona por la que se está propagando uno de estos haces, en su camino a la Tierra. Como el efecto de la onda es contraer y expandir el espacio, el haz sufrirá un pequeñísimo retraso o adelanto relativo a lo que hubiera ocurrido sin onda gravitatoria. Medio período de la onda más tarde, lo que era retraso se transformará en adelanto. Lo que quiere decir que si a lo largo de varios años registramos con suficiente precisión los tiempos de llegada de los pulsos, el posible paso de una onda gravitatoria de ese período por la región del espacio por la que ha viajado el haz deberá manifestarse en pequeñísimas variaciones en los tiempos de llegada registrados de los pulsos, alternando a lo largo del tiempo fases con ligerísimos retrasos con otras de ligerísimos adelantos sobre los tiempos de llegada 'previsibles' (que serían los que habríamos registrado de no haber habido onda gravitatoria). Este método

requiere un programa de observaciones mantenido a plazo de varios años, y es factible directamente desde la Tierra con radiotelescopios.

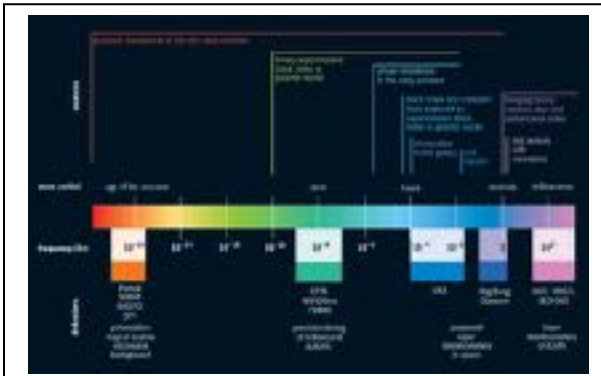


Fig.6. El espectro de la radiación gravitatoria, con las fuentes esperables a diferentes frecuencias y los instrumentos para su detección. Fuente: Presentación 'Status of the search for Gravitational Waves, Alan Weinstein, LIGO Scientific Collaboration, LIGO-G0900681.

Como se conocen bastantes pulsars de milisegundo, una mejora obvia es seguir a todo un conjunto de pulsars y extraer la información no solo de las observaciones de los retrasos y adelantos en la llegada de los pulsos de cada uno, sino de sus correlaciones en función de su separación angular. Como siempre, observaciones múltiples permiten reducir las incertidumbres observacionales, e identificar las amplitudes del fondo de ondas gravitatorias de esas frecuencias, que son las producidas por las fusiones de los agujeros negros supermasivos de los núcleos

galácticos; de este mapa se podría obtener información valiosísima sobre la formación de las galaxias o sobre sus fusiones. Actualmente hay tres proyectos en curso realizando estas medidas.

La astronomía y cosmología multimensajero.

Comparada con la fusión de agujeros negros, que emite solo ondas gravitatorias, la fusión de dos estrellas de neutrones emite además mucha radiación electromagnética generada en la sopa de materia ordinaria o ionizada que se expulsa en el momento de la fusión. Que eso es así realmente lo sabemos desde la detección del colapso de dos estrellas de neutrones GW170817 en agosto de 2017, que se pudo observar en ondas gravitatorias e inmediatamente después en todo el rango 'electromagnético'.

Pasar de una observación exclusivamente en el visible a poder conjugarla con la información obtenida en radio, microondas, infrarrojo, rayos X y rayos gamma fue un avance fundamental en la astronomía del S. XX. Ahora tenemos ante nosotros la posibilidad de complementar las observaciones 'electromagnéticas' con la información proporcionada por la astronomía 'gravitatoria'. Esta doble conjunción ya tiene un nombre: astronomía multimensajero. Las ondas gravitatorias transportan información imposible de obtener por medios electromagnéticos: por ejemplo, hemos visto ya que los agujeros negros, invisibles directamente por medios ópticos, pueden 'sentirse' a través de las ondas gravitatorias que un sistema binario de agujeros negros emite.

La información combinada que podemos obtener conjugando los canales electromagnético y gravitatorio es extraordinaria. Por ejemplo, se sabe ahora que, además de en la explosiones de supernova, la mayor parte de los

elementos químicos de número atómico mayor que el del hierro, como el oro, se producen precisamente en estos procesos de fusión de estrellas de neutrones. Y que estos procesos son el origen de algunos de los antes misteriosos brotes de rayos gamma. Seguramente esto solo es el comienzo de lo que podremos aprender así.

Las sirenas estándar.

Una buena muestra de la potencialidad de esta astronomía multimensajero es la estimación del parámetro de Hubble usando los datos de la observación de la onda gravitatoria GW170817 producida por la fusión de dos estrellas de neutrones. Esa estimación consigue la misma precisión que las realizadas 'ópticamente' alcanzaron en los 1990, tras 60 años de refinamientos. Cuando este método disponga de varias observaciones multimensajero semejantes, se podrá usar para obtener una determinación 'absoluta' de las distancias cosmológicas que no dependa de las delicadas calibraciones requeridas en los métodos actuales de determinación de esas distancias, agrupados en lo que se llama la escalera cósmica.

La idea básica fue propuesta por B. Schutz en los 1990s, y en su forma original consiste en observar sistemas que emitan en gravitatorio y en electromagnético, como el colapso de un sistema binario de estrellas de neutrones. De las señales gravitatorias conocemos la amplitud y la frecuencia de la onda gravitatoria recibida y sus ritmos de crecimiento en los ciclos finales antes del colapso. Nótese que no podemos concluir que la frecuencia observada sea la frecuencia original de emisión, debido al desplazamiento cosmológico de frecuencias, que no puede medirse con métodos puramente gravitatorios (recordemos que no hay 'rayas' gravitatorias). Pero de la observación electromagnética sí que podemos medir ese desplazamiento al rojo, que es el mismo para ambos tipos de ondas. Con esa información la teoría determina la amplitud inicial de las ondas en el momento de su emisión, que comparadas con la amplitud de la onda gravitatoria recibida, nos dan *directamente* la distancia a la fuente.

Por oposición al nombre de *candelas estándar* que se aplican a los métodos de la escalera cósmica que consisten en comparar luminosidades, a estos nuevos estándares se les ha bautizado en 2005 como *sirenas estándar*. Conserve ese término en mente, ya que sin duda dará bastante que hablar en un futuro próximo.

Las ondas gravitatorias primordiales.

Hay una barrera infranqueable para poder observar el Universo temprano mediante 'luz': el fondo cósmico de microondas. Descubierta accidentalmente en 1964, es literalmente un registro de cómo era el Universo con unos 380.000 años de edad. Aquella radiación, que se emitió más o menos en el visible, la podemos observar hoy en la región del espectro de microondas, muy desplazada hacia el rojo por la expansión del Universo. El estudio detallado de

esa radiación, que desde los 1990 se lleva haciendo con varios satélites específicamente dedicados, es uno de los tres pilares de la Cosmología actual. ¿Porqué no podemos 'ver' más atrás? Porque antes el Universo era (como una niebla intensa es para la luz) y solo entonces se hizo transparente, como lo es ahora, a la radiación electromagnética.

En contraste, el Universo siempre fue transparente a las ondas gravitatorias. Y en los primeros instantes, tras el Big Bang, con las altísimas densidades de materia y de energía en movimiento se debió emitir una gran cantidad de radiación gravitatoria, que, al igual que el fondo cósmico de microondas, se estará propagando desde entonces. Lo que significa que, en principio, no hay ninguna barrera temporal para poder 'observar' los primeros instantes del Universo mediante ondas gravitatorias.

La detección de esas ondas gravitatorias primordiales abriría unas posibilidades insospechadas. ¿Seremos capaces de conseguirlo? Nadie lo sabe, pero que pudiéramos hacerlo sería una fascinante revolución, mayor que la promovida por el descubrimiento del fondo cósmico de microondas.

A esa esperanza se refieren las palabras del principal galardonado con el Nobel de Física 2017 por la detección directa de las ondas gravitatorias, Rainer Weiss, con las que quiero cerrar esta rápida excursión cultural:

“There are calculations that indicate that the very earliest instants of the universe, right after the universe gets born, there is an enormous amount of background radiation of gravitational waves generated. That would be one of the most fascinating things man could [see] because it will tell you very much how the universe starts [...] At some point, not with the detectors we have now, we hope to be able to look at the beginnings of the universe.”



Fig.7. Dos de los ganadores del Nobel de Física 2017: Rainer Weiss, izquierda, Kip. S. Thorne, derecha, en el acto (Febrero de 2016) de presentación de la primera detección de ondas gravitatorias, hecha cuatro meses antes. Foto: Andrew Harnik, FILE / AP Photo / Gtres.

Referencias:

- [1] Thorne, K.S., Agujeros Negros y Tiempo Curvo, Drakontos, 2010 (publicado originalmente en inglés en 1994). (Libro general que describe el contexto de los agujeros negros y comenta muchos detalles históricos de primera mano).
- [2] Emparan, R., Iluminando el lado oscuro del Universo: Agujeros Negros, ondas gravitatorias y otras melodías de Einstein, Ariel, 2018. (Un libro reciente y excelente, por un autor español).
- [3] Thorne, K.S., Gravitational Waves, in Proceedings of the Snowmass'94 Summer Study on Particle and Nuclear Astrophysics and Cosmology, eds. E.W. Kolb and R. Peccei (World Scientific, Singapore, 1995), pp. 160-84; also published in Particle Physics, Astrophysics & Cosmology, Proceedings of the SLAC Summer Institute on Particle Physics, eds. Jennifer Chan & Lilian DePorcel (SLAC-Report-484, Stanford Linear Accelerator Center, Stanford, California, 1996). <http://xxx.lanl.gov/abs/gr-qc/9506086>
- [4] Barish, Barry C. and Weiss, Rainer, LIGO and the detection of gravitational waves, Physics today 52, 10, 44 (1999).
- [5] Abbott et al., Observation of Gravitational Waves from a binary Black Hole merger, Phys. Rev. Lett, 116, 061102 (2016). (Artículo original con el primer anuncio de detección, "muy técnico", pero es informativa la lista completa de autores y de sus instituciones, que ocupa casi tanto como el propio artículo).
- [6] González, G y Sintés, A.M., Ondas gravitacionales: mensajeras del Universo, Revista Española de Física, 29, 4, 2015.
- [7] Sintés, A.M., 14 de Septiembre de 2015: el albor de la astronomía de ondas gravitatorias, Revista Española de Física, 30, 2, 2016.
- [8] Sintés, A.M, Sorazu, B, La observación de ondas gravitatorias con LIGO, Investigación y Ciencia, Feb. 2017.
- [9] Castelvecchi, Davide. Here come the waves, Nature, 556, 12 April 2018.
- [10] Rothman, T. The secret history of Gravitational waves, American Scientist, March-April 2018
- [11] Sormani, C., Introduction, Notices of the AMS, 64, 7, 685; Denson Hill and Nurowski, P., How the green light was given for Gravitational Wave Search, Ibidem, 686, Bieri, L, Garfinkle, D and Yunes, N, Gravitational waves and Their Mathematics, Ibidem 693. (Aspectos más matemáticos excelentemente discutidos en este artículo doble publicado en las Notices of the American Mathematical Society bajo el título común The Mathematics of Gravitational Waves).
- [12] Nobel Committee Report for the Nobel Prize in Physics 2017. Popular Science Background, Cosmic chirps, www.kwa.se
- [13] Nobel Committee Report for the Nobel Prize in Physics 2017. Scientific Background, The laser interferometric gravitational-wave observatory and the first direct detection of gravitational waves, www.kwa.se
- (Resúmenes publicados por el Comité Nobel con ocasión de la concesión del Nobel de Física 2017: hay uno 'popular', accesible al mismo nivel que el presente artículo, y otro más 'avanzado', que es realmente excelente).
- [14] Sitios Web: Prácticamente todos los grandes proyectos tienen excelentes páginas web, con gran cantidad de información. LIGO: www.ligo.org , LISA: <https://www.elisascience.org/>

Mariano Santander Navarro

Doctor en Ciencias Físicas.

Catedrático del Departamento de Física Teórica, Atómica y Óptica,
Universidad de Valladolid.

Capítulo 25

El más allá del Modelo Estándar: futuros pasos en física de partículas.

¿Entendemos la física de partículas? ¡Por supuesto! Tenemos hasta un modelo. ¡EL! modelo: el Modelo Estándar de física de partículas. Esta es la primera respuesta que nos encontramos en muchos sitios y de muchas personas. Argumentos no faltan, desde luego, pues sabemos mucho de la física de partículas y este modelo es una de las teorías más exitosas de la historia. Pero... ¿sabemos lo suficiente? No voy a ser yo quien responda a esta pregunta, prefiero que lo hagas tú. Yo solo quiero contarte algunas de las ideas que se discuten a día de hoy y que podrían ser importantes en el futuro.

Sin embargo, antes de empezar déjame aclarar dos cosas. Primero, es imposible comprender el presente o imaginarse el futuro de nada sin conocer antes su historia. Nuestro pasado define quiénes somos y señala hacia dónde podría ir nuestro futuro. Por eso te recomiendo que antes de seguir leas el capítulo “El Modelo Estándar de la Física de Partículas”, de Claudia García en este mismo libro.

Segundo, hablar del futuro es muy difícil. Vamos a hablar de cosas que no se entienden hoy, pero que nos gustaría entender mañana. Como te podrás imaginar, es muy probable que parte -¡o todo!- de lo que digamos aquí resulte estar mal dentro de unos años. ¡No pasa nada! No por miedo a equivocarnos deja de ser divertido hablar de lo desconocido. Y si nos equivocamos, que lo haremos, ya se encargarán los historiadores del futuro de corregirnos.

Dicho esto, estamos preparados para hablar de algunos de los grandes pasos que podríamos dar para entender mejor la física de lo muy pequeño. Lo vamos a hacer sin ningún orden en particular y de la manera más independiente posible, de forma que puedas saltar directamente al tema que más te interese. Habrá temas que se queden fuera —como la gravedad, que la dejamos para el capítulo de Enrique Álvarez—, pero si te gustara saber más de otros temas no tengas miedo en preguntárnoslo. Empecemos.

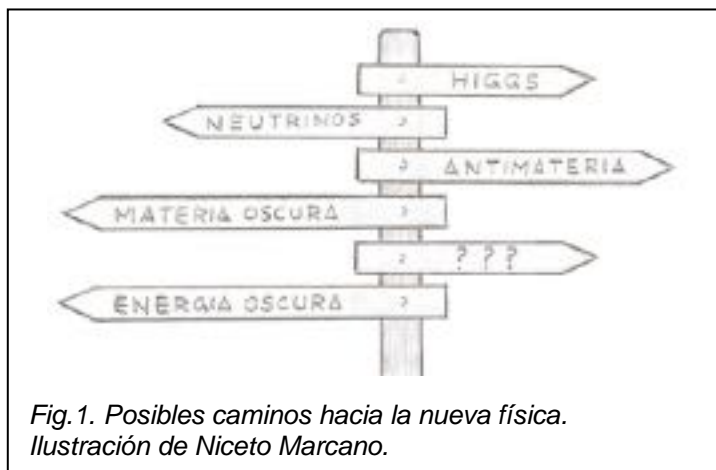


Fig. 1. Posibles caminos hacia la nueva física.
Ilustración de Niceto Marcano.

¿Higgs, eres tú?

El descubrimiento del bosón de Higgs en el LHC supuso encontrar la pieza perdida del Modelo Estándar y, como cuando se coloca la última piedra de un arco, le dio la consistencia que toda gran teoría necesita. Pero el Modelo Estándar no predice solo su existencia, sino también muchas de sus propiedades. Es hora de ponerlas a prueba.

El LHC ya ha conseguido comprobar algunas de las propiedades del bosón de Higgs — o simplemente Higgs, para los amigos —, como su carga y espín, o sus interacciones con algunas de las partículas más pesadas. Esto último es importante para ver que entendemos bien las cosas. Fíjate que decimos que la masa de una partícula depende de su interacción con el Higgs, así que los experimentos deberían ser capaces de ver que la intensidad con la que se relacionan con el Higgs está relacionada con su masa. Las de mayor



masa *se hablan* más fuerte con el Higgs, a grito pelado vamos, por lo que el LHC ya ha conseguido medir este efecto. Las más ligeras, en cambio, solo cuchichean con el Higgs, así que es muy difícil que nos enteremos de lo que dicen en el LHC, o incluso en otros experimentos que vendrán después. Pero bueno, con lo que ha visto el LHC hasta ahora parece que todo marcha a pedir del Modelo Estándar.

Pero todo esto del Higgs genera una pregunta inmediata. Si la masa de las partículas viene de su interacción con el Higgs, ¿de dónde viene la masa del propio Higgs? No es fácil contestarlo. La respuesta fácil sería algo como ‘yo *no he sido, estaba así cuando yo llegué*’. La respuesta real es más compleja, pero el mensaje es parecido. Resulta que el Higgs es distinto a las demás partículas del Modelo Estándar. No es un fermión que hace de *ladrillo*, pero tampoco es el bosón mensajero de ninguna fuerza. Es un bosón, sí, pero sin espín. Es el más soso de todos si quieres. Tan soso que nadie se ha molestado en prohibirle que tenga masa, como sí han hecho con el resto de partículas. El Higgs podía tener la masa que quisiera y, de hecho, hasta que apareció en el LHC no sabíamos que había elegido ser 125 veces más pesado que el protón. Por qué este valor y no otro es algo que no sabemos. De hecho, nos podríamos esperar algo mucho más grande. Pero la cosa es que saber este número era vital para el Modelo Estándar porque con él la teoría nos dice el resto de sus propiedades, como la velocidad con la que se desintegra o cómo de intensa es su interacción consigo mismo.

Sí, como lo lees, los bosones de Higgs socializan entre ellos y lo mejor es que no sabemos cómo. Bueno, miento, el Modelo Estándar sí que lo sabe. Nos dice la frecuencia con la que se juntan tres o hasta cuatro bosones de Higgs. Grupos más grandes ya no, porque se le va de las manos. Claro que estas son solo las predicciones teóricas, hace falta medirlo en los experimentos y, por desgracia, es muy difícil que el LHC consiga hacerlo. Es probable que tengamos que esperar a experimentos futuros, por lo que podemos decir que a

día de hoy no tenemos ni idea de cómo se comportan los bosones de Higgs en sus reuniones.

Llegados a este punto es probable que pienses que no sabemos nada del Higgs. Es la única partícula fundamental sin espín y no sabemos por qué. Y podrían tener una secta bosónica secreta montada entre ellos y nosotros aquí sin enterarnos. Bueno, razón no te falta, pero piensa que hace solo siete años que sabemos de su existencia, así que todavía nos queda mucho que aprender.

Sin embargo, es precisamente éste el motivo que lleva a mucha gente a estudiar en detalle todo lo que rodea al Higgs. Muchos intentan explicar por qué tiene la masa que tiene. ¿Te has fijado que se parece mucho a la masa de los bosones W y Z ? ¿Será casualidad? Hay a quien lo que le mosquea es que sea la única partícula fundamental sin espín y que piensa que si hay una, debería haber más — si hay tres familias de materia, ¿por qué iba a haber solo un Higgs?—. O puede que simplemente el Higgs no sea tan fundamental como dice ser, sino que esté formado de otras partículas que desconocemos. También hay quien piensa que podría estar relacionado con alguna nueva fuerza, como ocurre con los otros bosones. Quién sabe.

Un mundo abierto de posibilidades que descubriremos al estudiar cómo interactúa el bosón de Higgs con el resto de partículas y, puede que sobre todo, consigo mismo. Porque es precisamente en todo esto que todavía no sabemos dónde podría estar escondido el gran paso hacia la nueva física.

Esos malditos neutrinos.

Se dice de los neutrinos que son unos bichos raros, aunque, como ocurre a menudo, confundimos raro con incomprendido. Solemos decir que son distintos a las demás partículas que conocemos — ¿y cuál no? —, pero su mayor diferencia puede que sea la dificultad para estudiarlos en los experimentos.

Todo sea dicho, los neutrinos no lo ponen fácil. Son como ninjas capaces de atravesar la Tierra entera sin dejar rastro alguno. Entonces, ¿cómo los estudiamos? Es más, ¿cómo sabemos que existen? La verdad es que responder a estas preguntas podría llevarnos un libro entero, así que mejor resumirlo en tres ingredientes básicos: ingenio, tamaño y paciencia.

Hace falta ingenio para diseñar una forma de detectarlos, para lo que es importante un buen conocimiento de física de partículas. Los neutrinos son invisibles para nosotros y nuestras herramientas, pero en ciertas condiciones pueden hacer aparecer otras partículas que sí vemos, como electrones o muones. Midiendo estas últimas, podemos deducir que por ahí ha pasado un neutrino. Por supuesto, también hace falta mucho ingenio para desarrollar el experimento y la tecnología necesaria para medir estas señales, que en algunos casos no son más que pequeños destellos de luz casi indetectables. El problema es que, aún con el mejor de los detectores, la probabilidad de *cazar* un neutrino sigue siendo muy pequeña. Por suerte, esto se puede compensar con tamaño y paciencia, como cuando compras muchos boletos de la lotería y

juegas muchas veces. Por esto, normalmente los experimentos de neutrinos tienen detectores enormes y funcionan durante muchos años, para aumentar sus probabilidades de detectar muchos neutrinos y poder estudiarlos.

Gracias a su naturaleza evasiva, durante mucho tiempo han conseguido mantener ocultas algunas de sus propiedades más importantes, como su masa. Según el Modelo Estándar los neutrinos no tienen masa, y así se pensaba hasta que los experimentos *SuperKamiokande* y *SNO* demostraron lo contrario. Esto ocurrió en los años 1998 y 2002, respectivamente, cuando consiguieron medir un efecto que solo puede ocurrir si los neutrinos tienen masa: las oscilaciones de neutrinos.

Las oscilaciones de neutrinos son una consecuencia de la mecánica cuántica, pero podemos entenderlo con un jamón y una monja. ¿Has probado alguna vez a decir *jamón* muchas veces y muy rápido? Tras un periodo un poco confuso, te encontrarás diciendo claramente *monja*, y si sigues al rato volverás a decir *jamón*. Oscilarás entre una palabra y otra pasando por un periodo en el que dices una combinación de las dos. Con los neutrinos pasa lo mismo, oscilan entre los tres tipos de neutrinos que existen. Por ejemplo, un neutrino puede producirse de tipo muónico, pero luego detectarse como tipo electrónico. Esto fue lo que midieron los experimentos de *SuperKamiokande* y *SNO* primero, y muchos otros después.



Durante los últimos años han sido muchos los experimentos que han ayudado a aprender sobre los neutrinos. Sabemos que a los tres tipos de neutrinos les gusta mezclarse entre ellos, todos con todos, y que tienen tres masas distintas, aunque todavía no sabemos de cuánta masa estamos hablando. Tampoco sabemos cuál de ellos es el más ligero y cuál el más pesado, ni si los neutrinos y antineutrinos se comportan igual. Es más, ni siquiera sabemos muy bien cuál es la diferencia entre neutrinos y antineutrinos. Vale, esto es muy raro, ¿cómo no vamos a saber diferenciar una partícula y su antipartícula?

Las diferencias entre partícula y antipartícula son sus cargas. Son iguales en todo pero con cargas, por ejemplo la eléctrica, opuestas. ¿Qué pasa entonces si una partícula no tiene carga? Este es el caso del neutrino, que no tiene carga eléctrica. ¿Quiere decir esto que el antineutrino es igual que el neutrino? La pregunta real es un poco más compleja porque existen más cargas que la eléctrica, pero la respuesta es igual de simple: no tenemos ni

idea. Los experimentos de oscilaciones de neutrinos no pueden responder a esta pregunta, pero existen otros experimentos que sí y, con un poco de suerte, nos lo dirán en un futuro cercano. A día de hoy, no sabemos si un antineutrino es distinto a un neutrino —llamado neutrino de *Dirac*— o si son lo mismo —neutrino de *Majorana*—. Esto último sería un bombazo en la física de partículas, puesto que todos los demás fermiones tienen partículas y antipartículas bien distintas, y haría que los neutrinos fuesen realmente especiales.

Lo más interesante es que nada de esto ocurriría si, como dice el Modelo Estándar, los neutrinos no tuviesen masa. ¿Quiere decir que esta especie de metamorfosis cuántica de los neutrinos nos está diciendo que el Modelo Estándar está mal? ¡Sí! y sin duda necesitamos una nueva teoría más completa. El problema es que no sabemos cuál... De hecho, a día de hoy se han propuesto decenas de teorías, cada una con sus partes buenas y malas. En muchas de ellas, se piensa que existen nuevos neutrinos que todavía están por descubrir. Se piensa incluso que estos podrían ser la materia oscura que buscamos, o los responsables de haber producido la diferencia necesaria entre materia y antimateria para que exista el universo pero no el antiuniverso. Claro que para demostrar que estas teorías son ciertas deberíamos encontrar estos nuevos neutrinos, y por ahora ningún experimento ha sido capaz de lograrlo.

Poco a poco vamos comprendiendo los neutrinos. Una vez más, ¡el secreto está en la masa! ¿Tendrán los neutrinos alguna relación con el bosón de Higgs? Tampoco lo sabemos todavía. Lo que sí sabemos es que descubrir el modelo detrás de la masa de los neutrinos será un gran paso para la física de partículas.

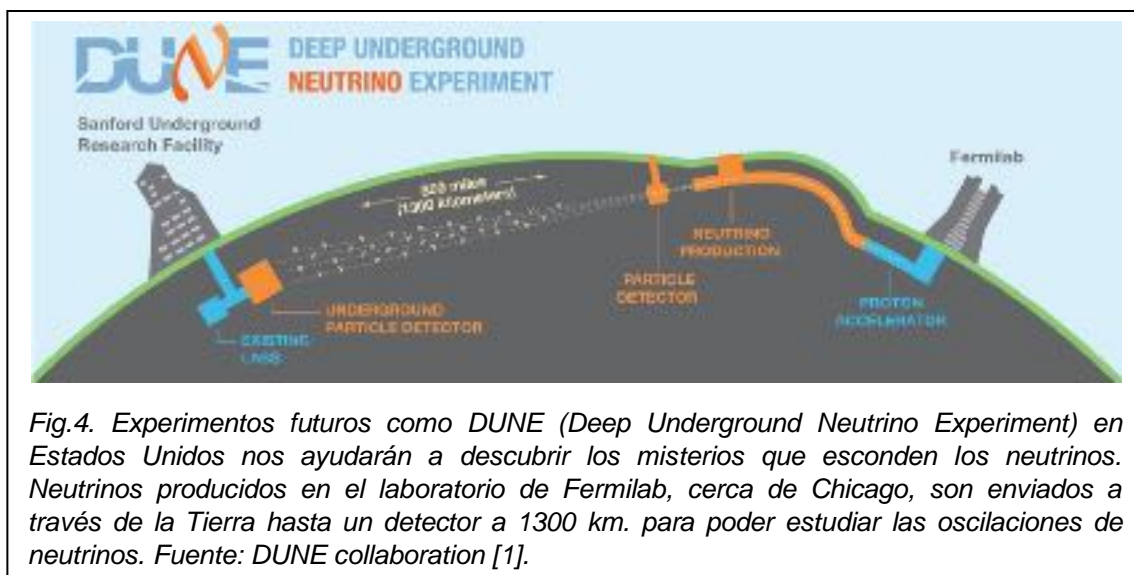


Fig.4. Experimentos futuros como DUNE (Deep Underground Neutrino Experiment) en Estados Unidos nos ayudarán a descubrir los misterios que esconden los neutrinos. Neutrinos producidos en el laboratorio de Fermilab, cerca de Chicago, son enviados a través de la Tierra hasta un detector a 1300 km. para poder estudiar las oscilaciones de neutrinos. Fuente: DUNE collaboration [1].

El antiuniverso perdido.

Todos perdemos cosas. ¿Quién no ha perdido alguna vez las llaves, la cartera o hasta el gato? Pero si miramos atrás, muy atrás, en la historia, nos

encontramos con la mayor pérdida de todas. ¡Hablamos de un universo entero! O más bien, de un antiuniverso.

Bueno, dicho así puede parecer que hablamos de una película de ciencia ficción, pero la realidad tampoco desmerece. Hablamos de toda la antimateria que se creó al principio de los tiempos y que ya no está. De la batalla entre materia y antimateria de la que salió victoriosa la materia y surgimos nosotros. Pudo ser al revés. Pudo no haber quedado nada tras la batalla.

La teoría del Big Bang nos explica cómo fue el principio de nuestra historia. Nos dice que poco después del Big Bang todo era una sopa muy caliente de luz, materia y antimateria, o mejor dicho, de las partículas y antipartículas que las componen. Como puede que sepas ya, al juntarse partícula y antipartícula se destruyen formando fotones de luz. Al mismo tiempo, los fotones se pueden juntar para crear parejas partícula-antipartícula, de modo que el proceso ocurre en las dos direcciones, como si de un partido de tenis se tratase. Según avanza el tiempo, el universo se enfría, haciendo este intercambio cada vez más difícil, como si la red de tenis se hiciese demasiado alta de superar. Fin del partido.

En realidad, lo que ocurrió se parece más a que, en un momento, los fotones se cansasen de jugar, de devolver la pelota en forma de pareja partícula-antipartícula. En el otro lado de la pista, en cambio, las partículas y antipartículas se siguen destruyendo mutuamente para generar fotones que lanzar, ignorantes de que al otro lado no hay nadie con quien jugar. El final de este partido solo puede ser uno: toda la materia acabará destruida junto con la antimateria, dejando un rastro de luz que ni en las mejores ferias. Entonces, ¿cómo es posible que existamos?

Dado que nosotros estamos hechos de materia, parece claro que en algún momento la balanza entre partículas y antipartículas se desniveló, de modo que al final de este partido de tenis, cuando la antimateria se abalanzó contra la materia, quedó un remanente de materia. En algún momento se crearon más partículas, o se perdieron más antipartículas, como prefieras. La diferencia fue de una partícula extra por cada mil millones de antipartículas. Parece un número ridículo, ¿verdad? Pero este 0.0000001% extra somos nosotros y el resto del universo que vemos.

Es increíble que una diferencia tan pequeña tenga unas consecuencias tan grandes, por lo que es inevitable preguntarse por su origen. Antes hemos dicho que la luz, los fotones, crean las partículas y antipartículas siempre en pareja, así que siempre tendremos el mismo número de ambas. ¿De dónde viene entonces esta diferencia?

Por suerte para nuestra existencia, no es luz todo lo que nos rodea. Existen otro tipo de interacciones, de fuerzas, entre las partículas, todas aquellas que describe el Modelo Estándar. Entre ellas está la interacción nuclear débil, que resulta que trata ligeramente distinto las partículas que las antipartículas. Esto quiere decir que, si el partido de tenis lo arbitrara la fuerza débil en vez de la electromagnética, el partido podría estar amañado a favor de la materia.

Este efecto de la interacción débil se ha estudiado bien entre los quarks en experimentos y se está estudiando ahora entre los neutrinos. Por desgracia, las predicciones del Modelo Estándar no son suficientes para explicar el exceso de materia del universo.

Este problema se conoce como *asimetría bariónica del universo*, en referencia a los bariones, como protones o neutrones, que forman la materia que vemos en el universo. A día de hoy seguimos sin saber qué originó esta diferencia entre materia y antimateria, pero existen toda clase de nuevas teorías que intentan explicarlo. Algunas de ellas, culpan a la materia oscura o a los neutrinos. Todas ellas son opciones interesantes, pero todavía no tenemos claro qué desequilibró la balanza, llevándose consigo un posible antiuniverso. Un gran misterio para resolver en el futuro.

El lado oscuro del universo.

Sabemos mucho de muy poco. Al menos eso es lo que nos dice el universo al mirar cómo se comporta. Todo lo que sabemos del Modelo Estándar y sus partículas no es más que el 5% de lo que forma el universo. El resto, un lado oscuro del que sabemos muy poco.

Al mirar al cielo, podemos observar planetas, estrellas y galaxias enteras. Podemos estudiar cómo se mueven, saber si se acercan o se alejan entre sí. También podemos ver cuánta materia hay en ellas. Esta materia que *vemos*, la materia visible, es a lo que nos referimos al hablar del Modelo Estándar y sus problemillas.

Sin embargo, el cielo nos cuenta mucho más. Por ejemplo, podemos estudiar cómo se mueven las estrellas dentro de sus galaxias. Según las leyes de la gravedad, el movimiento de una estrella depende de la cantidad de materia que hay en la galaxia, ya que es la que se encarga de atraerla y atraparla es su órbita. Pero como somos científicos, no nos fiamos de lo que dicen las leyes sin comprobarlo. Para ello, podemos medir la velocidad de las estrellas, compararla con la materia que vemos en su galaxia y comprobar que la ley de la gravedad se cumple. Esta medida se ha hecho en muchas galaxias y... ¡no sale bien! Las estrellas se mueven como si hubiese mucha más materia de la que vemos. ¿Quiere decir esto que la ley de la gravedad está mal?, ¿o es que hay algo más que no vemos?

Para responder a esta pregunta podemos seguir mirando al cielo. Empezamos por comprobar que las leyes de la gravedad funcionan en otros casos... parece que sí. ¿Será entonces que hay más materia que no podemos ver? Pero, ¿cómo vamos a estudiar algo que no vemos?

Una posibilidad en estos casos es simular lo que pasaría si existiese. Usando ordenadores muy potentes, podemos pedirles que sigan las leyes de la gravedad desde el principio del universo hasta ahora, y que nos digan cómo hubiese sido el universo si solo hubiese la materia que vemos. El resultado no es bueno. La gravedad hace que la materia se atraiga entre sí y se acumule en las galaxias que vemos hoy en día, pero solo con la materia que vemos la

gravedad no es lo suficientemente potente para que esto ocurra. ¿Y si metemos un poco de esta otra materia que no vemos? A más materia, mayor será la gravedad, podría funcionar. Y así es, funciona. De hecho, las simulaciones dicen que este nuevo tipo de materia podría haber tenido un papel clave en la creación de las galaxias y demás formaciones que vemos hoy en día.

Esta materia que no podemos ver recibe el original nombre de *materia oscura*. La razón para no verla es simplemente que no emite ni reacciona con la luz, o si lo hace es tan poco que no podemos verlo. Por suerte, su parte materia le hace sentir la gravedad y nos permite aprender un poco sobre ella. Por ejemplo, mirando choques de galaxias — ¡sí, las galaxias chocan! — aprendemos que la materia oscura no es muy sociable, que interactúa muy poco con el resto de materia de su entorno. Se parece un poco a los neutrinos, poco sociables y que no emiten luz. ¿Será que la materia oscura son neutrinos?

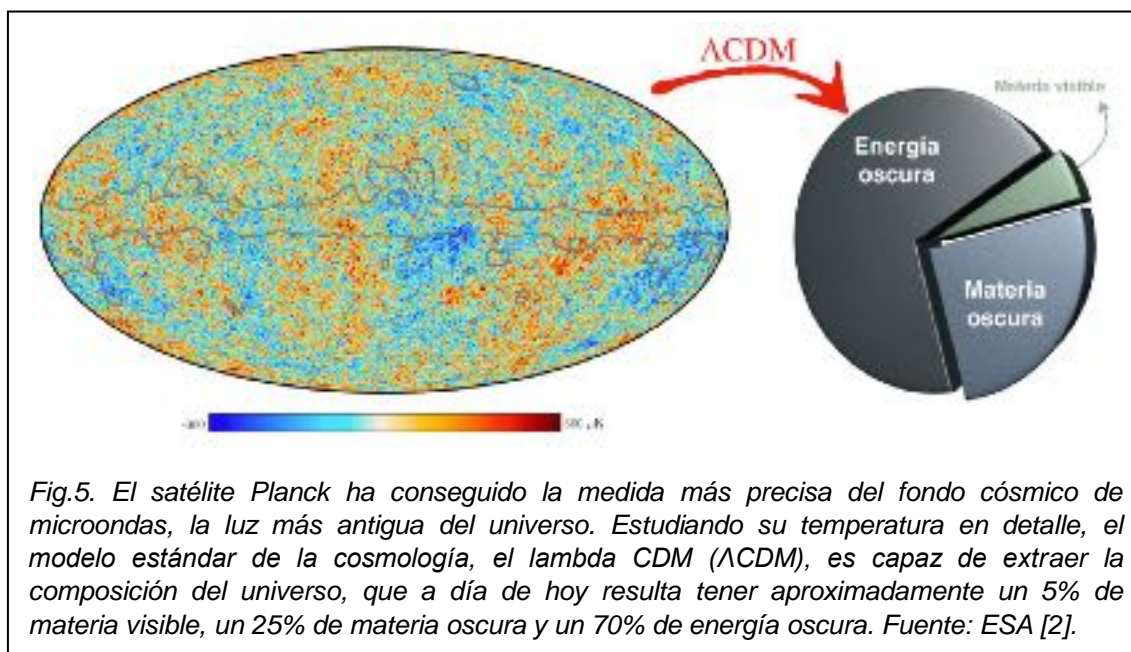
Los neutrinos son las únicas partículas del Modelo Estándar que podrían llegar a ser la materia oscura. Por desgracia, su masa es demasiado pequeña y esto impide que ayuden lo suficiente en la formación de las galaxias. Sin embargo, si existiesen nuevos neutrinos más pesados sí que podrían. Este es un ejemplo de cómo un modelo nuevo podría matar dos pájaros de un tiro, añadiendo nuevas partículas que estarían detrás de la masa de los neutrinos y serían la materia oscura que buscamos. No lo voy a negar, este tipo de soluciones conjuntas son muy interesantes y bonitas — en la escala de la física teórica —, sin embargo a día de hoy no sabemos qué demonios es la materia oscura.

En realidad, por saber no sabemos seguro si la materia oscura está formada por una — o varias — partícula fundamental que no hayamos descubierto todavía, aunque esta es la hipótesis más común. La esperanza de los rastreadores de materia oscura es entonces que estas nuevas partículas hablen, aunque sea muy bajito, con alguna de las partículas que sí podemos ver del Modelo Estándar. Los experimentos de materia oscura llevan varios años tras este rastro y, aunque no hayan conseguido encontrar al culpable todavía, la esperanza es lo último que se pierde.

No cabe duda que descubrir la naturaleza de la materia oscura sería un gran paso para la física, ya que estaríamos hablando de entender aproximadamente un 25% del universo, que sumado al 5% de la materia visible, hacen un total de... un mísero 30%. ¡Tanto lío para esto! ¿Qué pasa con el resto? El 70% restante es algo aún más oscuro y desconcertante... es energía oscura.

Para entenderlo, volvamos al cielo e imaginémosnos cómo se ha movido toda la materia que se creó tras el Big Bang. Al principio, con el impulso inicial, toda la materia tendería a alejarse entre sí. Pero claro, por otro lado tenemos la gravedad que intentará atraerla, así que hará de freno primero y pondrá rumbo de colisión después, como cuando lanzamos una piedra al aire y pensamos que apartarse no es buena idea. Aplicado al universo, nos deja dos opciones: o bien el universo todavía sigue expandiéndose, producto de la inercia inicial

pero frenado poco a poco por la gravedad; o bien la gravedad lo ha frenado ya y está empezando a colapsar. Esto se ha medido observando cómo se alejan algunas supernovas, que son muy útiles para estos estudios, y el resultado es que el universo se está expandiendo... ¡pero cada vez más deprisa!



Es como si algo estuviese empujando las galaxias entre sí, y es muy raro porque la gravedad es siempre atractiva. ¿Cómo puede ser? ¿Existe la *antigravedad* o algo así? En cierto modo, se podría decir que sí. Resulta que en las ecuaciones que propuso Einstein para explicar la gravedad puede existir una parte que produzca este efecto de antigravedad que explicaría la expansión acelerada del universo que vemos. Este bicho matemático se conoce como *constante cosmológica* y fue, en sus propias palabras, “el mayor error” de Einstein.

Al igual que nosotros, Einstein se imaginó el mismo destino para el universo, un final triste que no le gustó, e introdujo esta nueva parte matemática en sus ecuaciones que compensase la atracción de la gravedad y dejase un universo estático, tranquilo, como a él le gustaba. Porque la tranquilidad no se valora lo suficiente. El propio Einstein reconoció su error al introducir este efecto basándose en sus propias creencias, pues no tenía motivos científicos para hacerlo. Años más tarde, al ver que el universo se expande cada vez más rápido, ha resultado que la constante cosmológica está presente. Los genios son genios hasta en sus equivocaciones.

La materia, tanto la visible como la oscura, genera efectos gravitatorios atractivos, por lo que este tipo de efecto expansivo se asocia a algo distinto, una especie de *energía oscura* que sigue reglas un tanto distintas. Se suele relacionar a la energía que tiene el propio vacío, por eso también se conoce como *energía del vacío*.

Un momento. Esto de que el vacío tenga energía nos suena de algo. Se parece un poco a lo que ocurre en el mecanismo de Higgs, en el que lo que

nosotros pensábamos que estaba vacío está en realidad lleno de *campo de Higgs*. ¿Será entonces que el mismo efecto que hace que las partículas fundamentales tengan masa es el encargado de que el universo se expanda de manera acelerada? Este cálculo se ha hecho y el resultado es tremendamente malo. Horrible. Se equivoca en decenas de órdenes de magnitud, por lo que podría considerarse una de las peores predicciones de la historia de la física. Nuestro gozo en un pozo, seguimos sin saber qué es la energía oscura.

Como ves, el lado oscuro domina el universo. Tanto la materia como la energía oscura son un misterio todavía. Entenderlas será un gran paso para entender de qué está hecho la mayor parte del universo.

El Modelo Estándar, de héroe a villano.

El Modelo Estándar funciona muy bien, salvo algunos pequeños cabos sueltos en los que estamos trabajando. Nos dice cómo funcionan las partículas que transmiten las tres fuerzas, cómo clasificar las partículas que forman la materia y cómo el Higgs se encarga de dar masa a todo. Y funciona tan bien que ha pasado de ser una de las mejores teorías de la física, a convertirse en la peor pesadilla de cualquiera que quiera construir una nueva teoría. Pero entonces, ¿es éste el final del viaje?

Si le preguntásemos esto a un o una investigadora en física teórica, difícilmente obtendríamos un sí por respuesta, como mucho pondría pose de superhéroe y nos diría algo como “*puede que llegue el día en que alcancemos el final, un punto donde nuestro conocimiento no pueda continuar y tengamos que rendirnos; pero hoy no será ese día, no mientras yo siga aquí preguntándome el porqué de las cosas*”.

Y es que éste es el papel de los físicos y físicas teóricas. Su misión es evitar que la física se detenga. Su enemigo, el Modelo Estándar de partículas y su incesante afán de explicarlo todo con éxito. Su mejor arma, una simple pregunta: *¿por qué?* Sé que puede sonar raro que lo importante sea la pregunta, algo que cualquiera puede hacer. Puede, incluso, que pienses que los mejores científicos son aquellos que responden las preguntas, pero estos son solo los más famosos. Los mejores son quienes encuentran la pregunta, pues no hay respuesta sin pregunta.

Ya para terminar, te propongo un juego. Pongámonos por un momento la capa de superhéroe y miremos al Modelo Estándar. Tienes barra libre para preguntarte todos los *porqués* que quieras, no te cortes. Yo te doy algunos ejemplos para arrancar:

-Hemos dicho que los *ladrillos* que forman la materia existen en tres *familias*, copias cada vez más pesadas. ¿Por qué tres y solo tres? ¿Hay algo mágico en este número? ¿Habrás más?

-La diferencia entre las familias es la masa, y la masa se la da el Higgs, pero ¿por qué tienen la masa que tienen y no otra? Tenemos algunas partículas

muy muy ligeras, como los neutrinos, y otras muy muy pesadas, como el quark top. ¿Por qué son tan distintas?

-¿Por qué en el Modelo Estándar hay tres fuerzas y por qué son como son? Quizá existen más fuerzas, quién sabe, con nuevas partículas y todo, que están ahí aunque nosotros no las conozcamos todavía. Como un nuevo continente esperando a ser *descubierto*.

-Hace más de cien años se dieron cuenta de que la electricidad y magnetismo eran dos caras de la misma moneda, el *electromagnetismo*. Años más tarde, resultó que el electromagnetismo era una parte de una fuerza más compleja, la fuerza *electrodébil*. ¿Por qué parar aquí? A lo mejor la fuerza electrodébil viene de una nueva fuerza. Tal vez la fuerza nuclear fuerte también. Puede que incluso de la misma fuerza, y no sean más que distintas caras de la misma moneda. ¿Te lo imaginas? Una única fuerza para dominarlas a todas.

Este tipo de preguntas son el pan de cada día de muchos teóricos y teóricas. Por darte un ejemplo, hay quien va en busca de la *teoría de gran unificación*, este santo grial que mencionábamos en el que todas las fuerzas vienen de una más *simple*. Pretenden entender la propia estructura del Modelo Estándar, por qué la materia viene en *packs indivisibles* de tres o por qué hay partículas tan ligeras y otras tan pesadas. Encontrar la explicación a estas preguntas es muy difícil, sobre todo porque no vale con dar una respuesta, hay que justificarla con predicciones que podamos demostrar en los experimentos. Todo un reto que hasta ahora ninguna teoría ha sido capaz de superar.

Por suerte, estos *superhéroes del porqué* tienen la ayuda de los experimentales, quienes han sido capaces de encontrar fisuras en la reluciente armadura del Modelo Estándar. Aquí hemos contado algunas de ellas, pero déjame insistir en que hablar del futuro es difícil, y que lo que hemos dicho aquí hoy puede que sea mentira mañana. Habrá problemas resueltos y nuevos problemas por resolver. Como suele decirse, *no serán todos los que están, pero seguro que no están todos los que serán*.

El tiempo — más bien el trabajo de investigadores e investigadoras — dirá si estas fisuras que hemos visto son lo suficientemente grandes como para romper la armadura y obligarnos a construir un nuevo modelo, o si necesitaremos encontrar más. Y para ello habrá que seguir haciéndose preguntas. Porque recuerda, un buen *¿por qué?* puede ser el mayor de los pasos hacia una nueva teoría.

Referencias:

[1] DUNE Collaboration: <http://www.dunescience.org>

[2] Misión PLANCK de la ESA: <https://www.cosmos.esa.int/web/planck>

Xabier Marcano

Doctor en Física Teórica.

Investigador postdoctoral en el *Laboratoire de Physique Théorique* de Orsay, Francia.

Capítulo 26

Biomatemática: la “evolución natural” de las matemáticas.

En el siglo XVII se cimentó la fructífera simbiosis científica entre la física y las matemáticas que ha posibilitado muchos de los grandes avances científicos y tecnológicos de la humanidad, como por ejemplo, el desarrollo en el siglo XX de la teoría de la relatividad que permite los cálculos para llegar a la Luna y más allá, para dominar la energía nuclear o, por citar algo más cotidiano, para desarrollar e implementar el sistema de posicionamiento global, conocido por GPS.

Dicha influencia recíproca llega a ser tan profunda que algunos grandes matemáticos, como por ejemplo V. I. Arnold (1937-2010), consideran que las matemáticas forman parte de la física y por consecuencia braman⁹² contra la tendencia, que apareció ya a mediados del siglo XX, de obviar la física en los estudios de matemáticas. Personalmente podría estar de acuerdo con muchas de sus consideraciones, pero, poco a poco, las diatribas de Arnold van perdiendo fuerza ya que la biología está comiendo terreno a la física en el papel de inspirador de las matemáticas y estoy seguro de que estimulará la creación de ámbitos matemáticos completamente nuevos. En ese sentido, tal y como dice Joel E. Cohen en su famoso ensayo⁹³, se puede afirmar que la biología será la nueva física de las matemáticas y las matemáticas el nuevo microscopio de la biología. De esa interacción entre matemáticas y biología surge la disciplina científica conocida como biomatemática o biología matemática de cuyo pasado, presente y futuro trataremos en este capítulo.

Definición de la biomatemática.

Hay muchas definiciones de la misma, casi todas coincidentes en lo esencial - la biología matemática o biomatemática es un área científica que estudia los procesos biológicos utilizando técnicas y modelos matemáticos-. Como muy acertadamente se precisa en la versión francesa⁹⁴ de Wikipedia, que traduzco a continuación:

⁹² web: <https://www.uni-muenster.de/Physik.TP/~munsteg/arnold.html>

⁹³ “Mathematics Is Biology’s Next Microscope, Only Better; Biology Is Mathematics’ Nest Physics, Only Better” Plos Biology 2004

⁹⁴ De façon précise les biomathématiques sont constituées par l'ensemble des méthodes et techniques mathématiques, numériques et informatiques qui permettent d'étudier et de modéliser les phénomènes et processus biologiques. Il s'agit donc bien d'une science fortement pluridisciplinaire que le mathématicien seul (ou le biologiste seul) est incapable de

La biomatemática es una ciencia altamente multidisciplinar que el matemático, informático o biólogo no puede desarrollar por sí solo. Para nacer y vivir, esta disciplina requiere equipos interdisciplinarios impulsados por el sentido de lo concreto. Lo abstracto (las matemáticas y sus teoremas) no son más que un medio para lograr una mejor comprensión de los fenómenos biológicos.

Delimitar el campo de actuación de la biomatemática es muy complicado ya que tiene interacciones con muchos otros campos, como biología de poblaciones, fisiología, genómica, farmacología, biología de sistemas, etc. De hecho, en este capítulo veremos varios ejemplos de ello.

Aunque los importantísimos y vertiginosos avances de estos primeros años del siglo XXI en las ciencias biológicas, propiciados fundamentalmente por la capacidad de secuenciación del genoma, estén dando impulso y relevancia a la biomatemática, esta disciplina no es una recién llegada al mundo de la ciencia, sino que lleva varios siglos entre nosotros.

Primeras evidencias de la utilidad de las matemáticas en biología.

La “Society for Mathematical Biology” considera⁹⁵ como nacimiento de la biomatemática la publicación en 1917 del libro “On Growth and Form” de D'Arcy Wentworth Thompson (1860-1948) y a su autor como el primer biomatemático, pero hay numerosos casos anteriores de aplicaciones de las matemáticas al estudio y descubrimiento de fenómenos biológicos.

Entre ellos podemos destacar como primer ejemplo un descubrimiento espectacular en el que las matemáticas aportan la prueba más convincente. Aconteció el 17 de abril de 1616 cuando William Harvey (1578-1657) presentó por primera vez a los miembros del Royal College su descubrimiento de que la sangre, bombeada por el corazón, circulaba por las venas y arterias y que no se generaba ininterrumpidamente en el hígado, a partir de los alimentos, como se había sostenido desde Galeno durante casi 1500 años. Uno de sus argumentos principales consistió en calcular el volumen de sangre expelido cada hora por el ventrículo izquierdo ~ 27 litros/hora y de ahí deducir que, si la sangre se generase en el hígado, la masa expulsada diariamente por una persona sería de 500Kg.

Ya en la segunda mitad del siglo XVIII, aparece lo que podríamos considerar el primer uso de las matemáticas en epidemiología cuando Daniel Bernoulli (1700-1782) desarrolla un modelo teórico⁹⁶ para expresar la proporción de individuos susceptibles de una infección endémica en términos

développer. Pour naître et vivre cette discipline exige des équipes interdisciplinaires mues par le sens du concret. L'abstrait n'est qu'un moyen pour parvenir à une meilleure compréhension des phénomènes biologiques. Les biomathématiques sont des débouchés tant pratiques que théoriques dans de nombreux domaines comme la biologie des populations, la physiologie, la génomique, la pharmacologie, etc. <https://fr.wikipedia.org/wiki/Biomath%C3%A9matique> (visitada el 16 de enero de 2019)

⁹⁵ <https://www.smb.org/history-of-the-society/>

⁹⁶ Bernoulli, D. Mém. Math. Phys. Acad. R. Sci. Paris 1– 45 (1766)

de la fuerza de la infección y la esperanza de vida. El principal objetivo⁹⁷ de Bernoulli era calcular la tabla de vida ajustada si se eliminara la viruela como causa de muerte.

Aportaciones matemáticas a la teoría evolutiva y genética de poblaciones.

A principios del siglo XVIII, por razones fundamentalmente económicas, surge el interés por el estudio de la demografía y en particular por la dinámica de las poblaciones humanas. En 1748, Leonhard Euler (1707-783) en su tratado "Introductio in analysis infinitorum" propone un modelo de crecimiento geométrico como modelo de crecimiento de la población humana. En 1798 Thomas Robert Malthus (1766-1834) advierte en su tratado "An essay on the principle of Population" de los obstáculos para que el crecimiento de la población humana fuese geométrico debido a que llegaría un momento en que la tierra no daría suficiente comida para alimentar a la población. Esa idea de Malthus dio lugar al desarrollo por Verhulst (1804-1849) en 1838 de la ecuación logística e influyó⁹⁸ poderosamente en Darwin (1809-1882) para llegar a la idea de la selección natural como principio para el origen de las especies, como así reconoce el propio Darwin en su autobiografía.

Matemáticas en el descubrimiento indirecto de la existencia de los genes.

Pero si algo destaca en la aplicación del método científico y las matemáticas (esta vez representadas por la estadística en su versión más elemental) en la ciencia biológica anterior al siglo XX, son los trabajos de Gregor Mendel (1822-1884) que marcan el nacimiento de la genética. Del análisis detallado de sus resultados experimentales, perfectamente diseñados y realizados, Mendel deduce tres leyes que constituyen la base teórica de la genética y que pueden ser comparados, por su importancia, a las leyes de Newton.

La conclusión de Mendel, totalmente contraria a las creencias de su tiempo, de que los caracteres hereditarios son discretos (genes) y no continuos, es una de las más importantes contribuciones de la Matemática a la biología. De hecho es una demostración de la capacidad y utilidad de las matemáticas como instrumento para deducir, a partir de resultados experimentales, propiedades

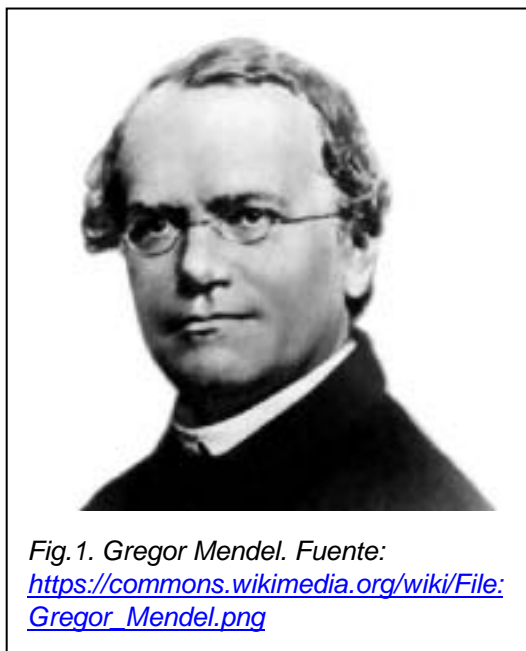


Fig.1. Gregor Mendel. Fuente: https://commons.wikimedia.org/wiki/File:Gregor_Mendel.png

⁹⁷ <https://www.nature.com/articles/35046270>

⁹⁸ También influyeron en la política demográfica China con resultados catastróficos.

de lo que no se puede observar directamente y, por lo tanto, generar nuevas hipótesis de trabajo que guíen la investigación y experimentación de biología. Esta capacidad de las matemáticas para ver lo que se esconde detrás de los datos está siendo, en estos comienzos del siglo XXI, crucial para el aprovechamiento de la ingente cantidad de datos que se están generando de forma continua y a ritmo acelerado en los numerosos laboratorios de investigación dispersos por el mundo. Se puede afirmar, sin temor a pecar de exagerados, que las matemáticas son el nuevo microscopio de la biología tal y como titulaba Joel E. Cohen en su ensayo anteriormente citado.

El siglo XX comienza con el redescubrimiento, en 1900, de los trabajos de Mendel y sus conclusiones sobre la transmisión de los caracteres hereditarios, que se pueden resumir en lo que hoy conocemos como las leyes de Mendel.



En 1922 Sir Ronald Aylmer Fisher (1890-1962), considerado como el padre de la estadística moderna, que desarrolló gran parte de su labor inspirado o buscando soluciones para problemas de índole genético, biológico y de la agricultura, publica su tratado "The Genetical Theory of Natural Selection". Con este tratado, en el que con un gran fundamento matemático se combinan las leyes de Mendel con la teoría de la selección natural de Charles Darwin, Fisher demuestra que el mendelismo valida el darwinismo.

Otros muchos matemáticos y biólogos siguieron desarrollando la idea y en 1942 Julian Huxley (1887-1975) en su tratado "Evolution: The Modern Synthesis" acuña el nombre de la teoría biológica "The Modern Synthesis" que dominó el pensamiento y la investigación en biología hasta el último cuarto del pasado siglo XX. En un cierto sentido, se puede ver un paralelismo entre la Síntesis Moderna y las leyes de Newton de la física. Actualmente se sabe que la teoría no es completamente correcta.

Matemáticas del desarrollo. Las ecuaciones de reacción difusión.

A nivel multicelular hay muchos campos en los que los modelos matemáticos están ayudando a la comprensión de los problemas biológicos. Cabe citar por ejemplo el caso de la oncología o del estudio del cerebro humano. Pero entre ellos me quiero fijar en un caso especial: el de la biología del desarrollo. Uno de los grandes desafíos de la biomatemática es proporcionar una descripción

matemática del desarrollo biológico. Un caso particular sería dar un modelo matemático del crecimiento de un embrión. Uno de los primeros intentos en este sentido se debe a A. Turing (1912-1954) que en su artículo “The Chemical Basis of Morphogenesis” publicado en 1950, propone un modelo que explica la formación de patrones en el crecimiento del embrión. Su modelo original consiste en suponer la existencia de dos especies moleculares, que llamó morfogenes, que estarían sometidas a procesos de difusión y reacción en el tejido embrionario lo que según él daría lugar a la aparición de patrones periódicos en dicho tejido. Turing es consciente de que el modelo está en estado muy primitivo y advierte al comienzo de su artículo que el modelo es una simplificación y una idealización y por lo tanto una falsificación. El modelo propuesto tuvo una influencia muy grande en numerosos campos científicos, tanto para explicar las formaciones de patrones en reacciones químicas, como en la formación de galaxias e incluso en la aparición de “puntos calientes” en la difusión del crimen en las ciudades. Sin embargo, se ha cuestionado mucho la validez del modelo en el campo para el que fue creado. Pero el tiempo le va dando la razón a Turing y cada vez aparecen más ejemplos de aplicación de las ecuaciones de reacción difusión para explicar patrones de desarrollo como el hecho de que tengamos cinco dedos, las rayas de los tigres o de las cebras. Es evidente que a medida que profundizamos en el conocimiento de los mecanismos moleculares podremos hacer modelos que se ajusten mejor, pero hay que reconocer una vez más la capacidad de A. Turing que, partiendo de una información muy poco precisa y escasa, fue capaz de vislumbrar los mecanismos que dirigen la formación de patrones en el desarrollo embrionario.



Fig.3. Alan Turing. Fuente: https://commons.wikimedia.org/wiki/File:Alan_Turing_Aged_16.jpg

Matemáticas para comprender la complejidad de la célula.

Si en la naturaleza muerta la unidad fundamental es el átomo, en la naturaleza viva (excluidos los virus que se encuentran en tierra de nadie) la unidad básica es la célula, ya que todo organismo vivo está constituido por una o más células. En el interior de toda célula está su genoma (molécula o moléculas de ADN que contienen todo su material genético).

La secuenciación (determinación de la sucesión de nucleótidos que forman las cadenas de ADN) del genoma humano, a principios de siglo XXI, cincuenta años después de la publicación del artículo en el que James Watson y Francis Crick –premiados con el Nobel en 1962– describían la estructura de

doble hélice del ADN, ha dado lugar a una inmensa revolución científica en el campo de la biología y las matemáticas han jugado y están jugando en ello un papel esencial.



Fig.4. Leonhard Euler. Fuente: https://commons.wikimedia.org/wiki/File:Leonhard_Euler.jpg

Para secuenciar el ADN hay que romper varias copias de la misma molécula aleatoriamente en millones de trozos que luego se secuencian en los secuenciadores automáticos. Para ensamblar dichos trozos se utilizan algoritmos muy complejos basados en teoría de grafos. El concepto de grafo, actualmente ubicuo en matemáticas y en muchas disciplinas científicas, fue introducido por L. Euler (1707-1783) unos 250 años antes de su uso en biología, para resolver un divertimento: el problema de los siete puentes de Königsberg. Poco a poco, los grafos han ido ganando presencia en matemáticas y actualmente es objeto de numerosos artículos de investigación y probablemente sean los objetos matemáticos más utilizados, sobre todo para intentar entender la estructura de los sistemas complejos. Volveremos a hablar un poco más de ellos en breve.

El genoma de una célula contiene, entre otras informaciones, su código genético, es decir la secuencia de todos sus genes, pero no llega con conocerlo para comprender el funcionamiento celular. La célula es un sistema muy complejo y dinámico en el que conviven numerosas moléculas químicas que están transformando e interaccionando entre ellas de manera continua. En el interior de una célula e incluso en su membrana hay millones de proteínas de diversos tipos que realizan la mayor parte de las funciones celulares. Por ello, el siguiente paso, una vez descifrado el genoma de un organismo, es conocer su proteoma, es decir, “el catálogo de sus proteínas”.

Uno de los organismos unicelulares más estudiados es la levadura (*Saccharomyces cerevisiae*). Su proteoma se estima que consta de unas 5.858 proteínas diferentes⁹⁹. El número de copias de cada una de esas proteínas presentes en la levadura en un momento determinado es muy variable y depende de muchos factores externos e internos, pero se estima¹⁰⁰ que en términos medios, una célula de levadura contiene $4,2 \times 10^7$ moléculas de proteínas.

⁹⁹ Saccharomyces Genome Database: <https://www.yeastgenome.org/>

¹⁰⁰ Cell Syst. 2018 Feb 28;6(2):192-205.e3. doi: 10.1016/j.cels.2017.12.004. Epub 2018 Jan 17. Unification of Protein Abundance Datasets Yields a Quantitative *Saccharomyces cerevisiae* Proteome. Ho B, Baryshnikova A, Brown GW. (<https://www.ncbi.nlm.nih.gov/pubmed/29361465>)

El proteoma humano es mucho más extenso. El análisis del genoma humano permite predecir, utilizando algoritmos basados en modelos de probabilidad, la existencia de unos 20.300 genes que codifican proteínas. A día de hoy (febrero 2019) hay evidencia¹⁰¹, porque se han encontrado, de 17.694 (87,2%) de ellas, por lo que todavía quedan unas 2.000 que se ocultan a los instrumentos, bien porque están mal predichas, bien porque su expresión está restringida a circunstancias excepcionales que no somos capaces de recrear en los laboratorios o bien porque sus propiedades biofísicas y químicas no son compatibles con la tecnología de detección existente. El número de copias de cada una de esas proteínas en los 230 tipos de células diferentes que hay en el cuerpo humano es muy variable y cambia continuamente en función de las circunstancias con las que se enfrenta la célula en todo momento, pero se estima que en una célula humana el número medio¹⁰² de proteínas es cercano a 10^9 .

Décadas de investigación en biología celular, biología molecular, bioquímica, biología estructural y biofísica, junto con el uso de técnicas bioinformáticas y grandes recursos informáticos, ha permitido crear grandes bases de datos, como por ejemplo UniProt¹⁰³, que contienen un compendio del conocimiento de dichas proteínas y sus funciones individuales. A este respecto y debido a que la función de una proteína depende de su estructura tridimensional, uno de los grandes problemas (físico, matemático y computacional), aún sin resolver, es conseguir un algoritmo de predicción de la estructura tridimensional de una proteína a partir de la lista de sus aminoácidos. Se trata de un gran problema de optimización global, que aparentemente la proteína resuelve mediante la solución de problemas más pequeños de optimización local¹⁰⁴.

Una célula no es solo un conjunto de genes y proteínas. La biología de Sistemas.

Si bien esa lista de proteínas proporciona un catálogo de los componentes individuales de una célula, no basta para entender el funcionamiento de una célula de la misma forma que la lista de las partes de un avión no nos aporta conocimiento sobre su funcionamiento. Necesitamos tener un plano detallado del avión para saber cómo se reúnen y trabajan conjuntamente todas esas partes. De manera análoga necesitamos dibujar un diagrama exhaustivo de redes genéticas y sus interacciones bioquímicas para entender cómo los genes y proteínas trabajan conjuntamente, porque las proteínas rara vez actúan solas. Las propiedades biológicas de una proteína dependen de su interacción con otras moléculas, en particular con otras proteínas. Las proteínas de una célula forman parte de diferentes “maquinarias moleculares” (pathways) que realizan

¹⁰¹ <https://www.nextprot.org>

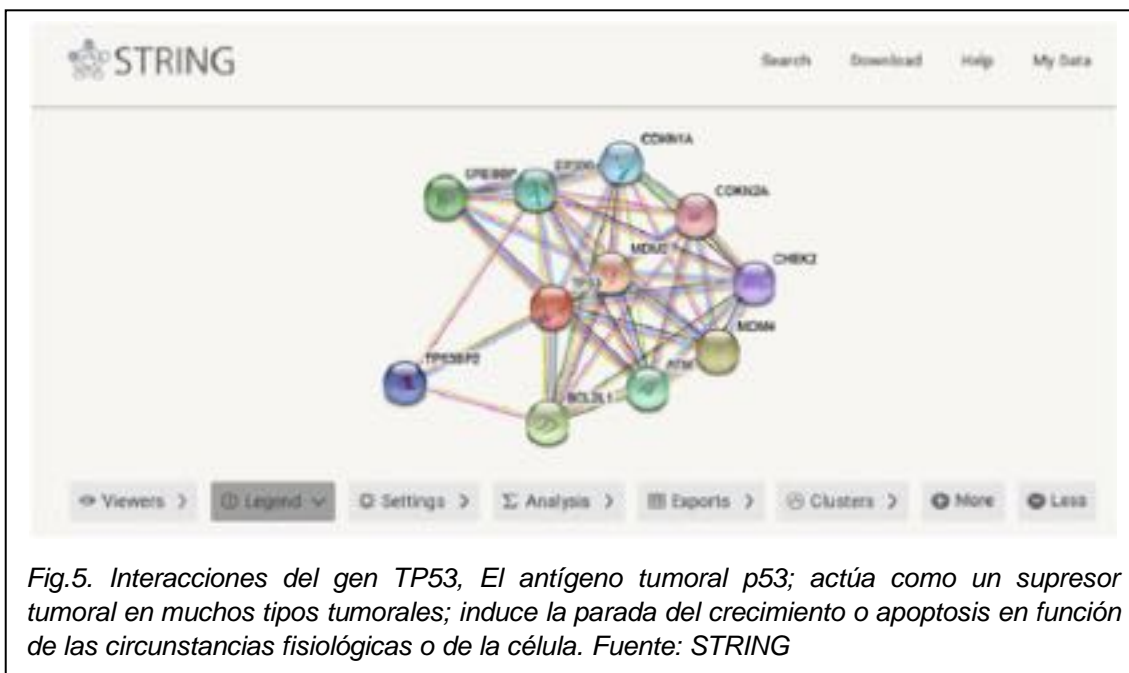
¹⁰² What is the total number of protein molecules per cell volume? A call to rethink some published values Ron Milo Bioessays. 2013. (<https://www.ncbi.nlm.nih.gov/pubmed/24114984>)

¹⁰³ UniProt: <https://www.uniprot.org/>

¹⁰⁴ Annu Rev Biophys. 2008 Jun 9; 37: 289–316. The Protein Folding Problem Ken A. Dill, S. Banu Ozkan, M. Scott Shell and Thomas R. Weikl. (<https://doi.org/10.1146/annurev.biophys.37.092707.153558>)

una o varias funciones. Los grafos (¡una vez más!) o “redes” son los objetos matemáticos que se están utilizando como modelos “primitivos” de esas máquinas moleculares. Se llama Interactoma a la red de interacciones entre las proteínas de un organismo.

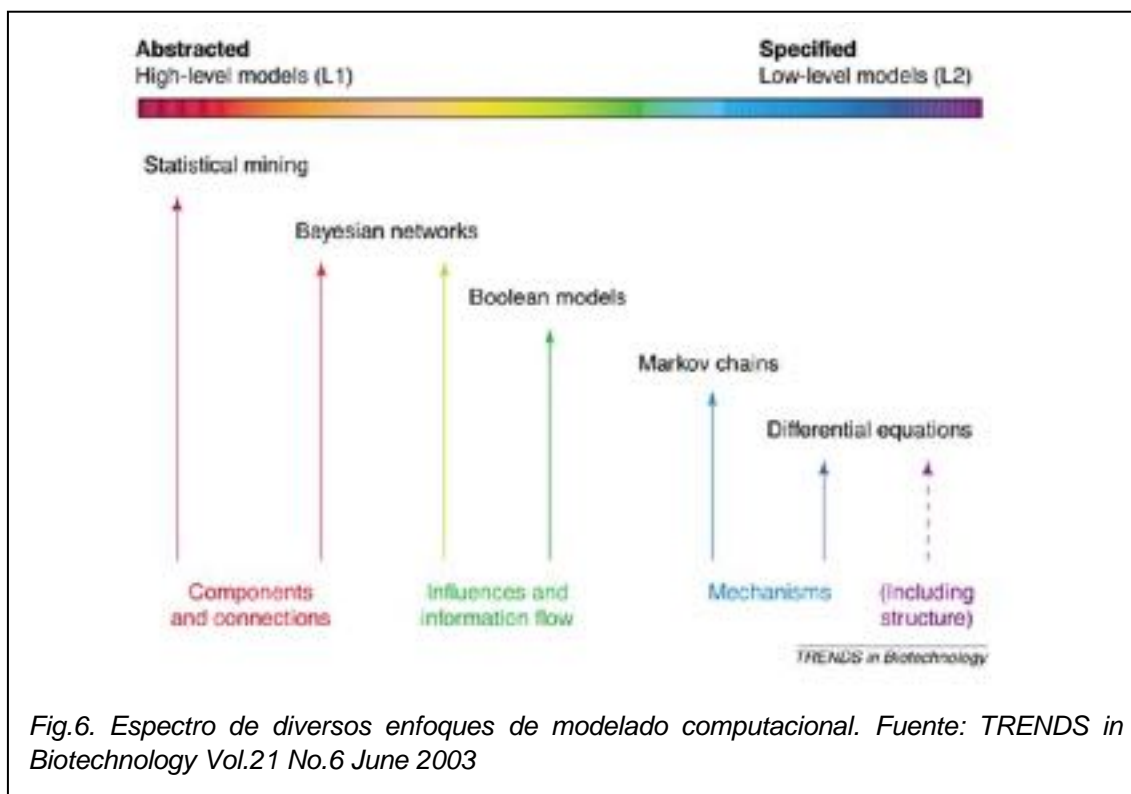
El estudio matemático de las propiedades del Interactoma, como por ejemplo la centralidad de sus vértices, la modularidad o la existencia de subredes pequeñas (motivos) sobrerrepresentados frente a un modelo aleatorio, está aportando información sobre los mecanismos moleculares que rigen la relación entre el genotipo “lista de genes de un organismo” y el fenotipo “expresión física de un organismo” además de generar hipótesis de trabajo para el desarrollo de nuevas investigaciones.



Pero el Interactoma no es más que un mapa de carreteras, su estructura nos puede ayudar a entender los procesos, pero no nos permite predecir ni controlar un sistema biológico. Por ello, a partir de pequeñas subredes del Interactoma, se están desarrollando modelos matemáticos, generalmente sistemas de ecuaciones diferenciales, que permitan entender el comportamiento de un conjunto pequeño de genes y su dinámica. Este tipo de modelización se enmarca en una disciplina emergente que se denomina biología de Sistemas.

Los instrumentos matemáticos que se utilizan para desarrollar los modelos en dicha disciplina van desde las redes booleanas a los sistemas de ecuaciones diferenciales estocásticas (ver figura anterior). El avance es lento, fruto de un trabajo colectivo que se desarrolla en laboratorios dispersos por el mundo, pero que tienen un objetivo común: hacer un modelo de todos los procesos biológicos que ocurren en una célula. Pero no podemos caer en la ingenuidad de pensar que los modelos van a ser perfectos. Como decía Turing, los modelos serán siempre simplificaciones de la realidad, es decir falsificaciones, pero poco a poco los iremos modificando y completando para

que al menos sean útiles. Por ejemplo, útiles para saber cómo intervenir, detener o controlar el crecimiento de un tumor.



La tarea es ingente, fuera del alcance de la tecnología y de las matemáticas que se conocen actualmente, probablemente utópico, pero alcanzar dicho objetivo es una de las grandes fuerzas impulsoras de la investigación en biología molecular. Los avances de este proceso cooperativo se pueden ver en varias bases de datos como por ejemplo, STRING¹⁰⁵ para el mapa de interacciones, KEGG¹⁰⁶ como base de datos de mapas de pequeñas redes llamadas "pathways" que realizan funciones específicas o BIOMODELS¹⁰⁷, una base de datos del EBI (European Bioinformatic Institute) que contiene información detallada sobre los modelos matemáticos que se van desarrollando.

¿Veremos algún día un modelo matemático completo de un organismo superior, y podremos simular en el ordenador la repercusión de la modificación de uno de sus genes o el efecto de un medicamento? ¿Podremos simular el efecto del intercambio de genes entre especies tal y como nos muestran en una de las películas de Spiderman? Sinceramente creo que sería posible a muy largo plazo, un siglo o dos, mientras tanto espero que los aprendices de brujo no se dediquen, como ya ha hecho algún investigador insensato, a la modificación de genes en humanos u otras especies superiores sin conocer las consecuencias de lo que hacen.

¹⁰⁵ STRING: <https://string-db.org/>

¹⁰⁶ KEGG: <https://www.genome.jp/kegg/pathway.html>

¹⁰⁷ BIOMODELS: <http://www.ebi.ac.uk/biomodels/>

$$\frac{dM}{dt} = \alpha_M \frac{1 + K_1(e^{-\mu\tau_M} A_{\tau_M})^n}{K + K_1(e^{-\mu\tau_M} A_{\tau_M})^n} + \Gamma_0 - \tilde{\gamma}_M M,$$

$$\frac{dB}{dt} = \alpha_B e^{-\mu\tau_B} M_{\tau_B} - \tilde{\gamma}_B B,$$

$$\frac{dA}{dt} = \alpha_A B \frac{L}{K_L + L} - \beta_A B \frac{A}{K_A + A} - \tilde{\gamma}_A A,$$

$$\frac{dL}{dt} = \alpha_L P \frac{L_c}{K_{L_c} + L_c} - \beta_{L_1} P \frac{L}{K_{L_1} + L} - \beta_{L_2} B \frac{L}{K_{L_2} + L} - \tilde{\gamma}_L L,$$

$$\frac{dP}{dt} = \alpha_P e^{-\mu(\tau_P + \tau_B)} M_{\tau_P + \tau_B} - \tilde{\gamma}_P P$$

Fig.7. Modelo del operon Lac, propuesto por Yildirim y Mackey en el año 2003.

Conclusiones.

La interacción entre la biología y las matemáticas no ha hecho más que empezar. La biología afronta una edad de oro que la convierte en la ciencia por excelencia de este siglo XXI y los venideros. La instrumentación científica que tienen a su disposición los investigadores del campo produce ingentes cantidades de información que tiene que ser analizada, clasificada y compartida a través de grandes bases de datos. Los modelos matemáticos que se desarrollan a partir de los datos sirven para generar nuevas hipótesis que se comprueban en los laboratorios cuyos resultados se reanalizan para ir afinando los modelos matemáticos. El objetivo final de toda esta investigación es ser capaces de predecir el desarrollo de un proceso biológico (como puede ser el crecimiento tumoral) y en aprender a controlarlo. Este trabajo de modelización implica a numerosas subdisciplinas de las matemáticas pero no puede ser realizado por matemáticos aislados en sus despachos y alejados del resto de los investigadores sino que tienen que ser fruto de un trabajo cooperativo e interdisciplinar. En mi opinión no es esperable la aparición de grandes teorías ni teoremas como así pasa en la física, sino que asistiremos a modelos cada vez más complejos y probablemente personalizados cuyas predicciones necesitarán de la computación en ordenadores cada vez más sofisticados. Probablemente la biomatemática aporte a la comunidad matemática problemas interesantes para ser estudiados teóricamente, pero ese no es el objetivo final de la misma.

Antonio Gómez Tato

Doctor en Matemáticas. Profesor Titular del departamento de Matemáticas de la Universidad de Santiago de Compostela.

Capítulo 27

Nueva ruta en la Neurociencia: manipulación de las bacterias intestinales que controlan el cerebro.

Estudiando el cerebro desde la barriga.

El dicho popular que afirma que en nuestras tripas se encuentra nuestro segundo cerebro, no podría ser más acertado hoy en día.

Los últimos avances de análisis de las secuencias bacterianas, su identificación por medio de poderosas herramientas bioinformáticas, y su manipulación con el objetivo de encontrar nuevos tratamientos para enfermedades neurológicas, representan como la neurociencia se está replanteando el futuro de los nuevos tratamientos.

En nuestros intestinos tenemos metros de longitud de pared intestinal inervada por neuronas. Tenemos incluso más neuronas en nuestra barriga que en nuestra cuerda espinal. Estas neuronas interactúan con las paredes intestinales donde se hospedan millones de especies de bacterias, lo que denominamos la flora intestinal o microbiota. La conexión del cerebro con los intestinos es una realidad, y sorprendentemente muy poco estudiada.

Tradicionalmente su estudio no presentaba mucho interés, tan solo se había descrito que estas bacterias contribuían a la digestión de los alimentos. Viven en simbiosis mutua, pacíficamente, en nuestro cuerpo, pero si las alteramos se revelan y pueden ser poderosas causantes de enfermedades.

En los últimos años ha surgido una revolución en el estudio estas bacterias, vinculándolas como una de las causas que originan una amplia variedad de enfermedades en el sistema nervioso central, como el autismo, enfermedades neurodegenerativas o incluso estados de ansiedad o depresión.

El artículo que estableció las bases del mapeado de la microbiota, el Proyecto de la Microbiota Humana¹⁰⁸, publicado por la prestigiosa revista *Nature* en 2012, describe un estudio a larga escala de los microbios que colonizan los humanos. Este análisis podría establecer las bases hacia futuros tratamientos para enfermedades neurológicas. Se está apostando por financiar investigaciones que estudien la relación de la microbiota intestinal y el cerebro, aportando millones de dólares a centros de investigación y científicos de estas disciplinas. Se especula que en la próxima década exista una explosión de avances científicos con nuevos tratamientos en esta línea.

¹⁰⁸ Proyecto de la Microbiota Humana (*Nature* 2012):
<https://www.nature.com/articles/nature11234>

Es muy importante que científicos de diversas disciplinas establezcan colaboraciones al respecto para entender los grandes hospedadores de nuestro cuerpo, y sobre todo para diseñar fármacos que puedan controlar sus acciones y curarnos de miles de patologías. Este es el momento cuando neurocientíficos, biólogos moleculares, microbiólogos o genetistas nos sentemos con bioinformáticos para analizar la secuencia genética de esta flora bacteriana y cómo se pueden encontrar las dianas adecuadas a las causas y curas de enfermedades que afectan al sistema nervioso, en otras palabras, cómo curar el cerebro estudiando lo que está ocurriendo en nuestras tripas.

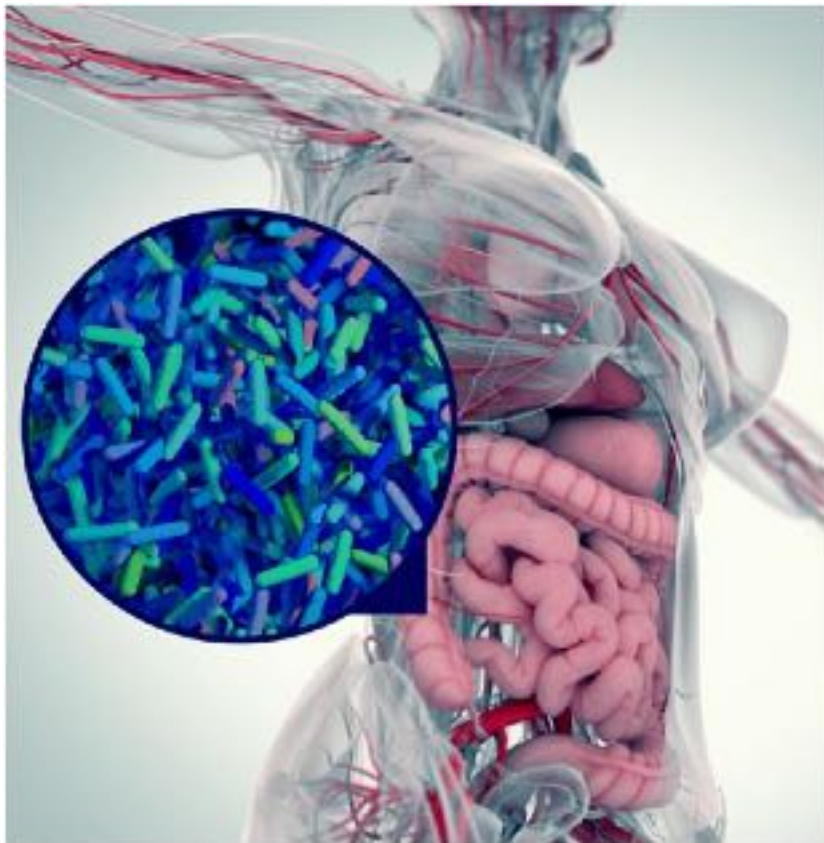


Fig.1. La flora o microbiota intestinal se compone de bacterias localizadas en el intestino y que conviven en una relación simbiótica.

Cuántas son y dónde están.

Cada ser humano es un entero ecosistema, donde convivimos en simbiosis con billones de bacterias de miles de especies diferentes. Sorprendentemente, estas bacterias han sido las gran olvidadas (aparte de intentar suprimirlas o eliminarlas), a pesar de constituir hasta un 90 % de nuestro cuerpo, quedando tan solo el 10 % restante a las células humanas. Se ha calculado que cada tres células de nuestro cuerpo, dos son bacterias y una es humana, aunque otros trabajos hablan de un ratio que se acerca a 1:1. Hasta tenemos su peso, en nuestro cuerpo se estima que tenemos alrededor de 200 g de bacterias. No

obstante, no se sabe todavía con certeza los números exactos, por lo que existen grandes controversias entre los científicos; de todos modos, todas las cifras propuestas son sorprendentes.

Si hablamos de los genes, es todavía más sorprendente. Solo el uno por ciento de nuestros genes es humano, el 99 por cien de los genes pertenecen a bacterias, levaduras, virus, y pequeños protozoos.

Las bacterias que habitan nuestro sistema digestivo, esta comunidad bacteriana denominada microbiota o flora intestinal, interactúan con el sistema endocrino, inmune y nervioso, afectando tanto a nuestro estado físico como mental, o influenciando en el desarrollo de muchas enfermedades. Desde el mismo momento de la gestación, la microbiota de la madre empieza a cambiar para favorecer el proceso. Todo lo que la madre come, determinará en cierta medida los cambios de la microbiota en el recién nacido. Si el parto es vaginal o por cesárea, o la lactancia materna, serán determinantes para adquirir un tipo de poblaciones bacterianas u otras. Estos habitantes microbianos son vitales para el desarrollo del sistema digestivo, metabólico o inmunitario del bebé. La microbiota adquirida en el momento del nacimiento es crucial y permanece en el desarrollo del individuo en su etapa adulta.

¿Cómo podemos alterar las bacterias intestinales?

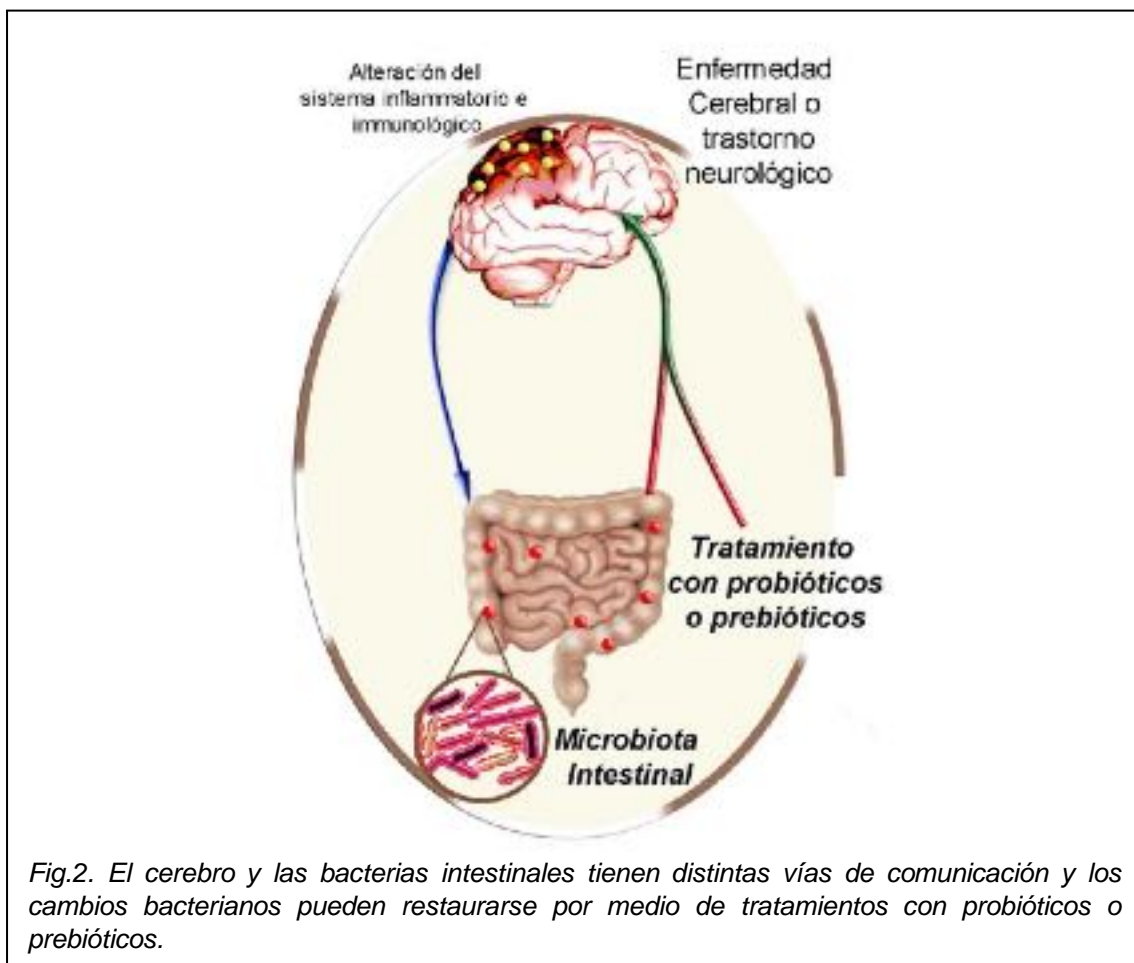
La microbiota o la flora intestinal a medida que crecemos se ve influenciada por la dieta, el entorno, el uso de medicamentos o antibióticos, lo que llamamos hábitos de vida. La dieta es crucial. Es la manipulación más sana que podemos hacer para mantener las colonias de bacterias beneficiosas para nuestro cuerpo. Añadiendo una dieta rica en fibras, estamos alimentando a nuestras bacterias con prebióticos que las harán resistentes a adversidades. La dieta tiene un efecto directo en la composición bacteriana, y es el mayor factor regulador de una salud intestinal adecuada.

Incluso a pacientes con depresión, se les recomendó seguir una dieta mediterránea con alta composición de polifenoles y ácidos omega 3 con propiedades antiinflamatorias, afectan a la composición bacteriana. Una dieta con comidas fermentadas que contienen probióticos naturales como el yogurt o el kéfir, vegetales cargados de fibra o buenos hábitos de ejercicio rutinario, tienen beneficios brutales en la “salud” de tu flora intestinal afectando al funcionamiento del organismo entero.

El abuso de antibióticos puede ser detrimental para muchas de las funciones fisiológicas que todavía desconocemos. Con abusivo uso de tratamientos de antibióticos se pueden perder grupos de bacterias fundamentales para nuestra salud. Aunque el problema se puede evitar repoblando estos grupos con diseñados probióticos de diseño que restauren la flora microbiana afectada.

Las bacterias y el cerebro, ¿cómo se comunican?

Las bacterias intestinales pueden profundamente afectar a la manera de funcionar nuestro cerebro. No se sabe qué especies o qué combinación de especies de bacterias pueden favorecer algunos procesos vitales que pueden verse alterados conduciendo así al desarrollo de enfermedades. Pero el efecto de la microbiota en nuestra salud va más allá. A pesar de décadas estudiando el cerebro como una identidad independiente, existe una evidente comunicación con el resto del cuerpo, y cómo no, también con el sistema gastrointestinal. Se ha demostrado que la microbiota libera metabolitos que pueden llegar al cerebro, afectando a muchas de sus funciones. Las vías de comunicación pueden ser por medio de pequeños neuropéptidos, metabolitos, o neurotransmisores que las bacterias producen, y que son rápidamente liberados a la circulación. Alternativamente pueden activar receptores neuronales en nuestros intestinos que por medio de fibras nerviosas, predominantemente a través del nervio vago, envían señales al cerebro y alternan mecanismos de acción. La administración a ratones de *Lactobacillus rhamnosus* favorecía la transcripción del ácido gamma-aminobutírico (GABA) a través del nervio vago, ya que al cortarse esta vía por medio de cirugía, no se observaron estos altos niveles de GABA. Las vías de modulación de esta comunicación son desconocidas y hoy en día no tenemos una noción precisa de los mecanismos de interacción.



Las bacterias que juegan con nuestras hormonas.

Sin duda alguna también las bacterias intestinales pueden manipular nuestras hormonas por medio de la liberación de pequeños neuropéptidos. Estas señales son las primeras en transmitirse al sistema nervioso y alterar nuestro sistema neuroendocrino, y por lo tanto nuestras emociones o conducta. Se ha encontrado que un tipo en particular de bacteria, el *Lactobacillus reuteri* puede aumentar los niveles de oxitocina, la hormona del amor y de los afectos. Una dieta alta en fibras está alimentando esta bacteria aportando un sabroso suplemento que la mantendrá a altos niveles en nuestro cuerpo. Otra de las hormonas cruciales para el estado de ánimo es la serotonina, conocida como hormona de la felicidad, que es producida mayoritariamente en el tracto gastrointestinal. Su regulación y producción dependen directamente de la composición de la flora bacteriana de cada persona.

La ansiedad y depresión se evitarán, modificando nuestras bacterias.

Hace pocos años era impensable afirmar que las bacterias que viven en el tracto digestivo pueden comunicarse con el cerebro e influenciar en nuestras emociones o inducir enfermedades. Hoy en día, existe una explosión de estudios científicos que lo demuestran.

Aunque la conexión "cerebro-tripas" se extiende a otras enfermedades, existen relaciones con los cambios en la microbiota y patologías neurológicas, como la ansiedad, depresión, autismo, o incluso la esquizofrenia. Se ha demostrado que ciertas condiciones de estrés inducen cambios en la diversidad y abundancia de la flora intestinal. Si lo miramos desde el punto de vista opuesto, si "reparamos" estos cambios bacterianos podríamos reducir en alguna manera los efectos del estrés.

Las llamadas bacterias malas o alteradas, son las que nos harán sentirnos antisociales o deprimidos. El hecho de tratar los problemas de salud mental con específicos probióticos va a ser una realidad en pocos años. Una vez que se realicen más estudios y se compruebe que tienen su efecto en grupos reducidos de pacientes, entraremos ante la nueva era de los psicobióticos en sustitución a la vieja farmacología para enfermedades mentales.

Enfermedades neurodegenerativas afectadas por bacterias.

Alteraciones en las bacterias intestinales también se han asociado al riesgo de padecer enfermedades neurodegenerativas. La enfermedad de Alzheimer, la enfermedad de Parkinson o la esclerosis lateral amiotrófica, cuyo origen genético representa un pequeño porcentaje, se desconoce la causa que las provoca en un 90 % de los pacientes. Cambios en la diversidad de las bacterias se han correlacionado con enfermedades cerebrales, e inflamación periférica en pacientes de Alzheimer. Se ha descubierto que los procesos que

desencadenan estas enfermedades podrían también estar dirigidos por proteínas producidas por las bacterias intestinales, como la proteína amiloidea. La síntesis de esta proteína y otros metabolitos que son absorbidos en la circulación sistémica pueden modular la función neuronal. Estos metabolitos pueden tener un efecto neuroprotectivo o neurotóxico, dependiendo del tipo de bacteria que los produce. Se ha descubierto cómo la exposición a proteínas bacterianas llamadas amiloides, que tienen una estructura similar a las proteínas del cerebro iniciadoras de muchas enfermedades neurodegenerativas, permiten la formación de agregados en el cerebro de otra proteína, la alfa-sinucleína. La proteína alfa-sinucleína es producida por las neuronas, causándole daños irreparables o incluso muerte neuronal, asociada a la patología de enfermedades neurodegenerativas, principalmente la enfermedad de Parkinson. Pero cómo estas proteínas producidas por las bacterias en el intestino pueden ser el origen de enfermedades neurodegenerativas sigue siendo una incógnita por resolver. Diversas hipótesis afirman que ciertas proteínas bacterianas que se producen en nuestros intestinos causan agregación de proteínas en el cerebro por medio de un mecanismo que permite empaquetar proteínas elásticas formando acumulaciones de proteínas que no se eliminan y aumentan en número, constituyendo el origen de la enfermedad. Estas proteínas bacterianas causan una bajada de las defensas inmunitarias en el sistema gastrointestinal lo que repercute aumentando la inflamación en el cerebro. Estudios en ratones expuestos al amiloide bacteriano desarrollaron agregados de alfa-sinucleína en el cerebro, mientras que ratones no expuestos no desarrollaron la enfermedad ni ninguna inflamación en el cerebro.

Aunque los mecanismos son desconocidos, se ha visto como después de la administración de una especie específica de bacteria en animales, eran capaces de recordar con más rapidez tareas motoras. Igualmente se demostró en personas que en estudios electrofisiológicos, se detectaron áreas cerebrales conectadas con mayor rapidez tras tomar un yogurt con *bifidobacterias*. Estudios en ratones aislados de patógenos y con una dieta esterilizada, tenían más neuronas en regiones que controlaban la memoria que los ratones convencionales, sugiriendo el papel de la microbiota en la inducción de la neurogénesis en el estado adulto.

La microbiota intestinal también envejece y se deteriora.

Los mecanismos por los cuales la microbiota cambia con la edad no son completamente entendidos. Ancianos que superan los cien años tienen una composición del microbioma muy particular y que difiere de otras personas 20 años más jóvenes. El estilo de vida, y en particular la dieta es determinante. La vejez viene acompañada por una reducción en la variedad de fibra consumida, y esto conlleva a una disminución en la diversidad de la microbiota, lo que puede ser dañino para la salud. Muchos estudios han demostrado que la alta diversidad de flora intestinal se asocia con un buen estado de salud.

Los metabolitos y componentes producidos por las bacterias de ratones envejecidos tienen propiedades pro-inflamatorias y asociadas al estrés

oxidativo y degeneración que la flora intestinal de ratones jóvenes. Las bacterias inducen efectos sistémicos en la producción de mediadores del metabolismo, la sensibilidad a la insulina o los estados inflamatorios. Por ejemplo, existen grupos de especies de bacterias asociadas a poblaciones envejecidas que, al trasplantarse a individuos jóvenes, conseguían avanzar en los procesos de envejecimiento.

Y otra vez más, ¿podremos en el futuro utilizar tratamientos de trasplantes de bacterias de personas jóvenes a personas mayores?, ¿conseguiremos de esta manera hacerlos rejuvenecer?

La población envejecida se incrementa exponencialmente y es requerida una mejor calidad de vida y un estado físico y fisiológico aceptable para llevar una calidad de vida aceptable. Aunque parezca ciencia ficción, el trasplante de flora intestinal no es una idea muy descabellada y posiblemente sea uno de los tratamientos revolucionarios de los próximos años. Los estudios actuales con animales pronto se trasladarán a humanos, y todo apunta que definitivamente la flora intestinal posiblemente sea la fuente de la eterna juventud.

Daños cerebrales o parálisis cambian la composición bacteriana, necesitamos reponerla.

También se ha visto cómo ratones con lesiones cerebrales o medulares se recuperan antes cuando se altera su microbiota, o cómo se produce una recuperación motora y neuropatológica. Existen estudios que relacionan las alteraciones en el microhábitat de las bacterias que habitan en el tracto gastrointestinal, llamado también *disbiosis*, con empeoramiento de las lesiones medulares, y por lo tanto con la recuperación motora. Estas alteraciones pueden estar causadas o bien por el ritmo de vida, el estrés, problemas gastrointestinales o incluso el uso descontrolado de antibióticos. Mediante un experimento muy curioso: administraron a ratones antibióticos potentes antes de inducirles una lesión medular, y posteriormente a un grupo de ratones se le administró una dieta rica en probióticos enriquecidos con ácido láctico y a otros ratones no. Curiosamente los ratones alimentados con probióticos respondieron mucho mejor a la fase de recuperación después de la lesión medular y su habilidad motora mejoró considerablemente, y por otra parte esto no ocurrió con los ratones con una dieta normal. Estos estudios sugieren que la alimentación con probióticos favorece la rápida recuperación después de lesiones medulares, reduciendo los procesos inflamatorios en el sistema nervioso y facilitando la función motora. Los nuevos tratamientos dirigidos a recuperar daños del sistema nervioso deberían mirar fuera de él. La microbiota que se aloja en nuestro intestino se comunica con el sistema nervioso directamente interaccionando con el sistema inmunitario o a través de fibras nerviosas, o indirectamente liberando metabolitos que pueden atravesar la barrera hematoencefálica.

Conclusiones.

La revolución del estudio de la microbiota intestinal y su comunicación con el cerebro es ya una realidad. En los próximos 20 años, con seguridad que va a ser una transformación en la manera que los científicos y los médicos estudian las enfermedades y sus distintos tratamientos. Queda mucho por explorar en este camino, pero las nuevas herramientas de análisis arrastran un futuro prometedor para que estos tratamientos individualizados se puedan aplicar con éxito. Otra pieza fundamental del puzle también se destapará, las causas y orígenes de muchas de las patologías que probablemente hayan estado escondidas en nuestras entrañas, descifrarlas y diagnosticarlas es la ruta venidera para el tratamiento de muchas de las enfermedades neurológicas.

Dra. Sonia Villapol

Neurocientífica y Profesora en el Departamento de Neurociencias de la Universidad de Weill Cornell (New York, EE.UU.) y en el Centro de Neuroregeneración (Houston, EE.UU.)

Capítulo 28

Exploración espacial: el gran viaje a nuestro Sistema Solar.

¿Aún no sabes dónde ir de vacaciones este verano? ¿Eres más de montaña que de playa? ¿De agua dulce o agua salada? ¿Te gustaría tomar el Sol en las regiones más cálidas de Sistema Solar, o por el contrario preferirías escalar sus montañas más altas, sus valles más profundos, o sus lagos perpetuamente congelados? Estamos a un paso de que estas preguntas se hagan realidad, y todo gracias a ¡la Exploración Espacial!

La imaginación es nuestra ventana al futuro. Gracias a los esfuerzos titánicos de las generaciones pasadas, a todo lo que hemos aprendido con ellas, y al gran despliegue actual de medios para el espacio (tanto público como privado), estamos llegando al punto álgido de la exploración espacial.

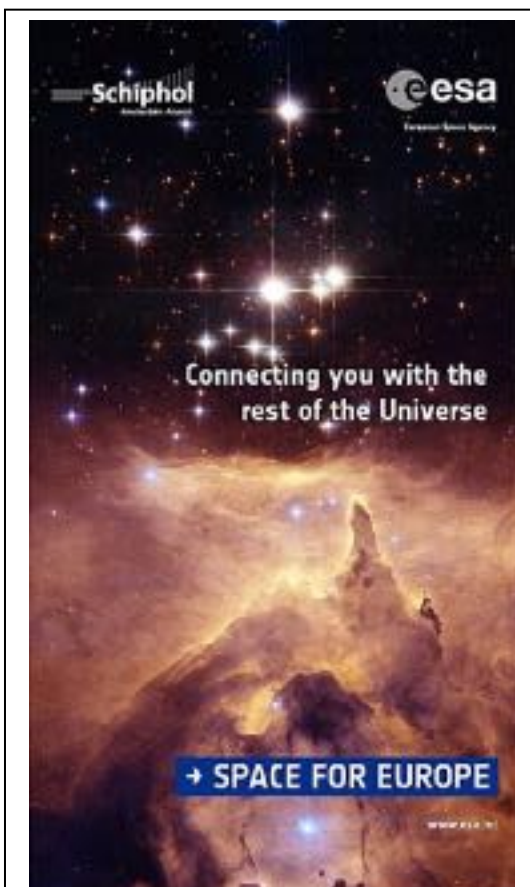


Fig.1. "Conectándote con el resto del Universo". Cartel promocional de la ESA en el aeropuerto Schiphol.

Ese punto en el que dentro de nada podremos mandar humanos a otros planetas, establecer colonias en ellos, e incluso encontrar formas de vida en otros mundos.

Regresar a la Luna, estudiar y viajar a Marte, alcanzar lugares más allá, son parte de las nuevas metas y retos que tenemos por delante. Las imágenes que ilustran este capítulo nos ponen en perspectiva de lo que otros mundos nos pueden llegar a ofrecer, y del porqué de la exploración espacial.

Pero... ¿Qué es la exploración espacial?

La exploración es una acción innata en todo ser humano. Desde el momento en el que nacemos sentimos curiosidad, nos sentimos atraídos por el mundo que nos rodea y lo exploramos, y lo más importante de todo, aprendemos de nuestra exploración. La exploración espacial consiste en exactamente lo mismo. Todo empieza con mirar al cielo

y sentir curiosidad de saber que se esconde más allá de nuestro planeta, y sentirnos atraídos por todo lo que no sabemos. De ahí surge nuestro deseo de explorarlo, y con ello, de aprender cómo otros mundos funcionan, cómo han evolucionado, o cómo interaccionan con el medio que los rodea. Explorar es nuestra principal herramienta para desafiar los límites del conocimiento, y además, nos ayuda enormemente a nuestro propio desarrollo científico, tecnológico, económico y social.

Actualmente, el objetivo principal es comprender el Sistema Solar en el que vivimos, entender su origen y evolución, por qué la vida comenzó en la Tierra, o si la vida podría haber comenzado o estar presente en otros lugares. En otras palabras, ¿por qué es así?, ¿qué fue antes, y qué sigue después?, ¿estamos solos? Además, aunque en este capítulo nos centremos solo en nuestro Sistema Solar, la exploración también abarca nuestra galaxia y más allá.

Gracias a la incansable exploración que comenzó hace 6 décadas, sólo en los últimos tres años hemos sido testigos de grandes hitos históricos, como el descubrimiento de las ondas gravitacionales producidas por la fusión de dos agujeros negros súper masivos, el descubrimiento de una Tierra gemela en la constelación del cisne, o el hallazgo del tan buscado depósito de agua líquida en Marte. De hecho, estamos a un paso de enviar humanos de vuelta a la Luna, de sentir posibles terremotos en Marte, de analizar muestras frescas de un asteroide, y quizás incluso de encontrar formas de vida en nuestro Sistema Solar. Todo esfuerzo tiene su recompensa, y ésta es un diamante en bruto.

Y... ¿qué me aporta a mí la exploración espacial como ser humano?

La exploración supone una gran inversión tanto de talento como de dinero. El pensamiento fácil que a muchos se les pasa rápidamente por la mente es ¿y para qué?, ¿no se puede usar todo ese dinero para mejorar nuestras vidas? Nuestra recomendación es sencilla. Sin moverte del sitio, mira a tú alrededor y piensa: ¿En qué aspectos ha mejorado mi calidad de vida gracias a la tecnología?, ¿sería fácil vivir sin todos los desarrollos tecnológicos que me rodean?

Aunque ahora no lo sepas, un muy alto porcentaje de la tecnología actual que nos rodea procede de diversas utilidades que han sido desarrolladas en el marco de la exploración espacial. Algunos ejemplos

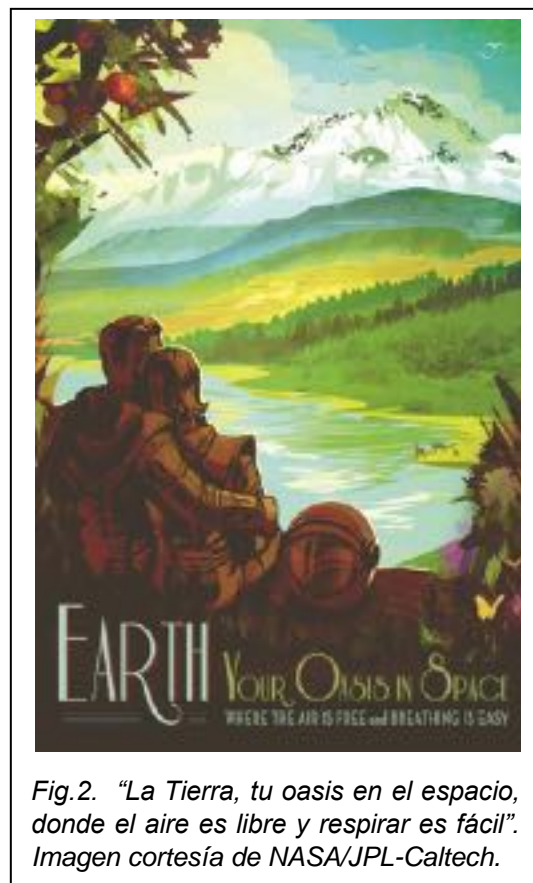


Fig.2. "La Tierra, tu oasis en el espacio, donde el aire es libre y respirar es fácil". Imagen cortesía de NASA/JPL-Caltech.

son los termómetros de oído, las nuevas prótesis para animales y humanos que son capaces de simular al máximo un miembro funcional, trajes ignífugos para bomberos, sistemas de radio que funcionan en condiciones de calor extremas, comunicaciones en tiempo real a largas distancias, los sistemas de navegación por satélite (como el GPS), el sistema internacional de salvamento, control de cosechas por satélite, aspiradoras sin cables, toda la comida deshidratada que nos rodea, la leche infantil, diseño de aerodinámica y niveles de estrés y vibraciones de vehículos, puentes y edificios que resisten terremotos, técnicas de fisioterapia por infrarrojos, válvulas cardíacas, instrumentación de muchísimas pruebas médicas, colchones con memoria, procesos de purificación del agua, detectores de humo, aislantes térmicos para diversos usos (como mantas, frigoríficos, motores, o ropa deportiva)... y un largo etcétera.

Viaje por el Sistema Solar interno: el Sol y los planetas terrestres.

Venus, Marte, la Luna... todos estos astros nos resultan muy familiares, y no solo porque hayamos estudiado sus nombres en el colegio, si no, porque al ser nuestros vecinos más cercanos, son fácilmente reconocibles en el cielo. ¿Quién no se ha quedado alguna vez embelesado con la Luna? ¿O quién no ha visto nunca el lucero del alba? Sin embargo... ¿Cuánto sabes en realidad sobre ellos?

La verdad es que nuestros vecinos planetarios tienen mucho que ofrecernos, y de los cuales podemos ayudarnos para entender la evolución que ha tenido nuestro Sistema Solar desde su origen. Sabemos que los planetas rocosos, Mercurio, Venus, la Tierra y Marte, se formaron al mismo tiempo, pero cada uno evolucionó de manera muy distinta. Por ejemplo, una vez Marte tuvo océanos de agua en su superficie y su atmósfera fue mucho más densa. Sin embargo, hoy en día Marte es un planeta árido, con una atmósfera tenue, y la única agua líquida que mantiene se encuentra en el Polo Sur del planeta en una cavidad subterránea como recientemente hemos sabido. ¿Qué le pasó a Marte? ¿Le podría pasar lo mismo a la Tierra? El estudio de estos cuerpos celestes, de cómo se comportan, de cómo se relacionan con el Sol, y de sus características principales nos da grandes pistas sobre su evolución y del por qué solo la Tierra ha sido capaz de proporcionar vida hasta la fecha (que sepamos).

El Sol, el corazón de nuestro Sistema Solar.

¿Qué decir del Sol? Vivimos gracias a él en todos los sentidos. Nos da luz, calor, y permite la vida en la Tierra, pero también, es el que ha dado lugar a la formación de nuestro Sistema Solar. Es sin duda el centro de nuestras vidas, pero en realidad, sabemos muy poco sobre él.

El Sol es una estrella típica catalogada como enana amarilla. Es una bola caliente de gases incandescentes que solo con su gravedad mantiene unido a todo el Sistema Solar. Las corrientes eléctricas en el Sol generan un campo magnético que se expande por todo el Sistema Solar en forma de corriente de gas cargado eléctricamente que sopla desde el Sol en todas



Fig.3. "¿Volando al Sol? ¡Nosotros sí!"
Cartel publicitario de Solar Orbiter por la
ESA y el aeropuerto Schiphol.

direcciones. Esto es lo que se conoce como viento solar. Un efecto muy visual de la interacción del viento solar con la Tierra son las auroras. Aunque el Sol es muy especial para nosotros, debemos recordar que hay miles de millones de estrellas como nuestro Sol dispersas en nuestra galaxia, la Vía Láctea.

Por razones obvias, la exploración in-situ del Sol ha estado largamente condicionada al desarrollo de una tecnología capaz de aguantar las altísimas temperaturas solares. Actualmente, tenemos muchos satélites que se dedican a la observación del Sol tomando imágenes y multitud de datos de la zona solar más externa, pero todos ellos lo hacen desde un punto cercano a la Tierra, a una distancia cautelosa. Sin embargo, tanto la agencia espacial americana (NASA) como la europea (ESA), han decidido dar un paso más y cambiar radicalmente el conocimiento que tenemos del corazón de nuestro Sistema Solar. La NASA envió en 2018 la sonda solar Parker para "tocar" el Sol en el sin duda sobrevuelo más cercano

realizado nunca a nuestra estrella, y la ESA hará lo propio en 2020 con Solar Orbiter, que también "tocará" el Sol y además visitará los polos solares que no son visibles desde la Tierra. Se espera que ambas naves aclaren grandes enigmas actuales que condicionan el medio ambiente de los demás planetas, como por ejemplo, el origen del campo magnético solar, el mecanismo que impulsa al viento solar, o el por qué y cómo de las erupciones solares de partículas energéticas (radiación) que luego se propagan por todo el Sistema Solar. ¡Pronto lo sabremos!

Mercurio y la misión europea BepiColombo.

Sobre Mercurio, la verdad es que poco sabemos. Es el planeta terrestre más pequeño, y menos explorado de nuestro Sistema Solar con diferencia. A pesar de estar tan cerca de la Tierra, es muy difícil de ver en el cielo porque se encuentra muy cerca del Sol y su luz lo oculta de nuestros ojos. Sin embargo, con los pocos datos que tenemos gracias a un sobrevuelo fugaz de la sonda de la NASA Mariner-10 en 1974, y de los cuatro años que la sonda MESSENGER de la NASA estuvo en órbita en Mercurio (2011-2015), sabemos que es un planeta fascinante, y en cierto modo, muy parecido a la Tierra desde un punto de vista geológico y magnético. La principal diferencia es que tiene una atmósfera extremadamente tenue, y en ese aspecto se parece más a nuestra Luna. Hasta ahora no disponíamos de la tecnología apropiada para acercarnos

tanto al Sol y mantener una nave a esas distancias, y por tanto, Mercurio había quedado un poco abandonado en la exploración espacial. Sin embargo, esto va a cambiar dentro de muy poco tiempo gracias a una nueva misión que actualmente se encuentra de camino a Mercurio y que fue lanzada en 2018. Se trata de la misión BepiColombo que consiste en dos naves que orbitarán este planeta gracias a un esfuerzo titánico de la agencia europea (ESA) y la japonesa (JAXA).

El viaje hasta Mercurio es muy largo, en concreto de 7 años. En realidad se podría llegar en tan solo unos meses, pero la gravedad del Sol atraería la nave de tal forma que no seríamos capaces de frenarla al llegar a Mercurio. Para evitarlo, es necesario realizar un viaje más largo, que además requiere de varios sobrevuelos a Venus con la intención de “frenar” la nave. A estos sobrevuelos a otros planetas se le conoce con el nombre de asistencias gravitatorias. Una vez que BepiColombo llegue a Mercurio allá por el 2025, tiene la ardua tarea de estudiar la formación y evolución de este planeta, lo que nos dará pistas sobre el nacimiento de nuestro Sistema Solar, así como de nuestro propio Sol. Estudiará en detalle su interacción tan cercana con el Sol, y como esto afecta a su superficie y al campo magnético del propio planeta. ¿Tiene Mercurio auroras como la Tierra, Marte, Júpiter o Saturno? También analizará la composición de su superficie y de su interior, si tiene un núcleo sólido y líquido como la Tierra, o si tiene alguna actividad geológica actualmente en su superficie. En otras palabras, ¿es un planeta volcánicamente activo?

Venus, nuestro gemelo malvado.

La exploración de Venus está en gran parte dominada por la motivación para estudiar un planeta que se considera como el gemelo de la Tierra, ya que ambos planetas se encuentran muy cerca, y son muy similares en masa y tamaño. Además, se sabe que Venus está activo geológicamente, con posibles volcanes potencialmente en erupción en la actualidad. Sin embargo, la pregunta principal es por qué a pesar de estas similitudes, ambos planetas tienen diferencias tan drásticas. Por ejemplo, Venus tiene un efecto invernadero extremo debido a una concentración muy elevada de dióxido de carbono, lo que provoca unas temperaturas extremadamente altas en su superficie, tan altas que pueden derretir el plomo. Además, un aspecto muy interesante de Venus es su muy lenta rotación, y además, en el sentido

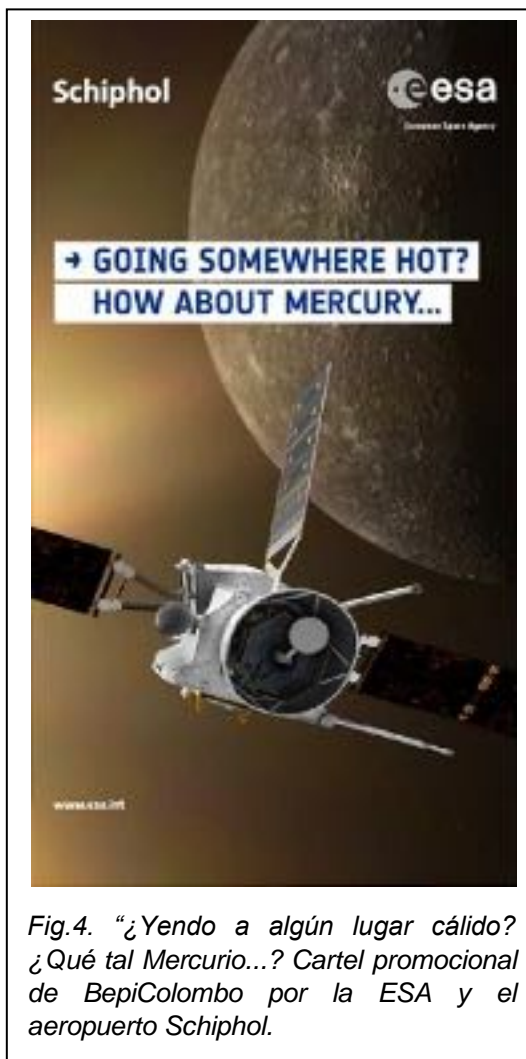


Fig.4. “¿Yendo a algún lugar cálido? ¿Qué tal Mercurio...? Cartel promocional de BepiColombo por la ESA y el aeropuerto Schiphol.

opuesto al resto de planetas, lo que pudiera haber tenido origen en un gran impacto en el pasado. Como curiosidad, un día en Venus es más largo que un año allí.

La capa de nubes de Venus no permite ver la superficie, por lo que todavía hay muchos misterios sobre la geología del planeta. Los resultados recientes de la misión europea Venus Express indican que algunos volcanes todavía pueden estar activos. De hecho, éste es el principal objetivo de la misión japonesa Akatsuki que actualmente se encuentra orbitando Venus desde 2015, así como de la futura misión india Shukrayaan-1 que planea lanzarse en 2023.

La Luna y Marte, primeros pasos para la exploración humana de otros planetas.

La vuelta de la exploración humana a la Luna es una realidad. Desde que la humanidad puso un pie en nuestro satélite natural en los años 60, la exploración robótica de la Luna ha sido incansable con multitud de países participando en su exploración. Sin embargo, ningún ser humano ha vuelto a poner un pie en su superficie desde 1972. Esto está por cambiar durante esta década puesto que varias agencias espaciales, entre ellas la NASA, ESA, Rusia, China, Japón e India están apostando, junto con compañías privadas, por la creación de una base lunar. Hasta ahora los planes son de misiones robóticas de exploración de la superficie de la Luna, así como con la creación



Fig.5. Imagen artística de la ESA sobre la futura base lunar.

de una plataforma orbital lunar. Algo así como una estación espacial internacional, para astronautas que vayan más allá de la Luna. Todos estos planes son ya una realidad. Como ejemplo cabe destacar el robot de superficie que China mandó a finales de 2018 a la cara oculta de la Luna, con semillas de plantas de algodón que llegaron incluso a germinar allí. Otro

ejemplo es la estrategia adoptada por la NASA para aprovechar la innovación de compañías espaciales estadounidenses para construir nuevos módulos de aterrizaje lunares. Toda esta exploración y desarrollo, a largo plazo nos brindará la experiencia necesaria para dar el próximo gran salto: misiones humanas a Marte, y más allá.

De hecho, uno de los principales objetivos que siempre ha tenido en mente la exploración espacial es llevar seres humanos a Marte. Todo comenzó en la época de la guerra fría cuando en plena carrera espacial tanto la Unión Soviética como los Estados Unidos se dieron cuenta de que los medios y

conocimiento que tenían a su alcance no eran suficientes en ese momento para mandar seres vivos a Marte. De ahí surgió que pensarán en crear diferentes estaciones espaciales habitadas de forma permanente para orbitar la Tierra y así prepararse para un día ir a Marte. Esto fue el germen que dio lugar años después a la Estación Espacial Internacional que cuenta con la presencia máxima de 6 astronautas, y en ella participan actualmente 5 agencias espaciales.

No obstante, la carrera espacial también fue el punto de salida de la exploración del planeta rojo de forma robótica. Desde 1960, más de 50 misiones han tenido por objeto este planeta. Pero poco a poco Marte se ha ganado a pulso el porqué de su larga exploración. Marte es un mundo fascinante que nos recuerda mucho a la Tierra, pero cuya evolución hacia una desertización total ha sido brutal. Muchos investigadores ven en Marte la imagen de lo que puede llegar a ser la Tierra si no respetamos a nuestro planeta. De hecho, es uno de los lugares del Sistema Solar donde muy probablemente podríamos dar respuesta a la pregunta sobre si hay o ha habido vida en el Sistema Solar. Sabemos que a lo largo de la historia, Marte ha sufrido cambios colosales en las condiciones de su superficie, donde todo indica que antes hubo océanos. Sabemos que estos cambios fueron producidos por bruscos cambios climáticos, así como cambios en la radiación solar que llegaba a Marte y gases de efecto invernadero.

Actualmente hay 8 misiones activas en Marte que estudian diferentes aspectos del planeta, desde posibles terremotos, hasta la continua pérdida de parte de su atmósfera en el espacio. Algunas de las futuras misiones incluyen la exploración de su luna mayor, Fobos, así como el estudio del análisis de muestras de este satélite de vuelta a la Tierra. Nos gustaría destacar la inminente misión de la ESA y Rusia, ExoMars, cuya primera parte se encuentra orbitando Marte desde 2016, y cuya segunda parte, un robot de superficie, será lanzado en 2020. Entre sus principales objetivos se encuentra el establecer si una vez existió la vida en Marte.



Fig.6. “Se buscan exploradores de Marte” Cartel promocional de la NASA para la exploración de Marte. Imagen cortesía de NASA/KSC.

Cometas y otros cuerpos menores, los grandes viajeros.

No podíamos olvidarnos de los grandes viajeros de nuestro Sistema Solar. Los cometas son, a grandes rasgos, bolas cósmicas de nieve de gases congelados,

rocas y polvo que orbitan alrededor del Sol. Cuando están congelados y lejos del Sol, suelen ser del tamaño de un pueblo pequeño. Pero cuando la órbita de un cometa se acerca al Sol atraído por su gravedad, se calienta y arroja polvo y gases que forman una cola que se extiende lejos del Sol por millones de kilómetros.

Algunos de los aspectos más importantes de los cometas desde el punto de vista exploratorio es que quizás son los cuerpos más antiguos y primitivos que aun preservan material de las primeras épocas de formación del Sistema Solar. Además, gracias a su composición rica en moléculas orgánicas, se piensa que en las primeras etapas de formación de la Tierra, cuando los impactos de cometas y asteroides eran mucho más probables que ahora, hubiera podido ser una forma por la cual la vida comenzara en nuestro planeta.



Fig.7. "Cita con un cometa". Cartel promocional de Rosetta por la ESA.

Desde los años 70, más de 10 misiones han tenido como finalidad sobrevolar diferentes cometas para su estudio con datos in-situ. Cabe destacar la primera imagen que se obtuvo del núcleo de un cometa, el cometa Halley, por la sonda europea Giotto en 1986. Pero sin duda alguna, si hay una misión que merece mención, esa es la misión europea Rosetta. Después de 10 años viajando por el Sistema Solar para llegar con la velocidad y precisión apropiada, Rosetta comenzó a orbitar el cometa 67P/Churyumov-Gerasimenko en agosto de 2014, y los dos juntos orbitaron alrededor del Sol por dos años. Más allá, Rosetta fue la primera misión que planteó y logró aterrizar con éxito otra sonda, Philae, en la propia superficie del cometa. Cabe destacar que gracias a Rosetta, nuestra imagen de los cometas ha cambiado radicalmente. Una de las principales

respuestas que esta misión nos ha dado es la confirmación de que el agua que el cometa 67P/Churyumov-Gerasimenko tiene es muy diferente al agua encontrada en la Tierra, así al que hay en otros cometas. Esto indica que el origen del agua que tenemos en la Tierra se podría deber a otros orígenes diversos. Nuevas misiones a cometas con diferentes niveles de actividad (que emitan más o menos niveles de polvo y gases) deberían ser una prioridad ya que estos cuerpos pequeños esconden información valiosísima sobre nuestros orígenes.

Además de cometas, nuestro Sistema Solar es rico en asteroides, masas rocosas que se encuentran orbitando alrededor del Sol. La mayor parte de ellos se encuentran en el cinturón de asteroides entre Marte y Júpiter, donde destacan algunos cuerpos más masivos como Vesta y Ceres. Estos mundos similares a planetas en pequeña escala, son supervivientes intactos de la

primera parte de la historia del Sistema Solar. Durante 11 años, la misión de la NASA Dawn estudió Vesta y Ceres hasta que en 2018 la sonda se quedó sin combustible. Uno de los principales resultados de la misión Dawn es el refuerzo de la idea de que los planetas enanos, y no solo las lunas heladas como veremos en la siguiente sección, podrían haber albergado océanos durante gran parte de su historia, y que algunos potencialmente aún lo hacen. Este campo de investigación está muy de moda por todas las oportunidades que ofrece. Actualmente tenemos una sonda japonesa, Hayabusa-2, que se encuentra orbitando el asteroide Ryugu de 1 km de diámetro y ha conseguido aterrizar dos pequeñas naves en su superficie con el objetivo de recoger muestras de material y traerlas a la Tierra para su análisis en 2020. También tenemos otra sonda americana, Osiris-REx, orbitando alrededor del asteroide de 500 m de diámetro Bennu. Pero no solo los asteroides son importantes desde un punto de vista científico, sino que pueden llegar a ser un riesgo potencial para nuestro planeta si sus órbitas se cruzan con la de la Tierra. Por eso, Europa planea mandar la misión Hera a un sistema de asteroides binario llamado Didymos, en el que ambos asteroides juntos son del tamaño de la pirámide de Guiza en Egipto. De hecho, Hera es una misión de defensa planetaria y complementa la misión DART de la NASA en la que esta nave impactará sobre uno de los asteroides de Didymos con el fin de desviar su órbita y probar así si esta tecnología es suficiente para desviar un asteroide en el caso de dirigirse a la Tierra.

Viaje por el Sistema Solar externo: los planetas gigantes.

La búsqueda de vida en el Sistema Solar siempre ha sido una de las grandes motivaciones para la investigación y la exploración. Por mucho tiempo creímos que Marte era el lugar más interesante para ello, pero ahora nos hemos dado cuenta de que las lunas de los planetas gigantes pudieran ser los lugares idóneos puesto que hay grandes evidencias de que bajo sus superficies congeladas hay grandes cantidades de agua, e incluso océanos. Por ello, llamamos a estos cuerpos “mundos oceánicos” y son actualmente una de las principales prioridades para la exploración.

Júpiter y la misión europea JUICE.

JUICE es una misión europea al planeta gigante Júpiter y sus 3 lunas heladas Europa, Calisto y Ganímedes. La misión será lanzada el 1 de junio de 2022 por un cohete Ariane 5, desde el centro espacial Kourou. Tenemos que ser pacientes, ya que el viaje de la Tierra a Júpiter durará 7 años y medio. Se necesitarán varios sobrevuelos a la Tierra, Venus y Marte para conseguir la velocidad suficiente para llegar a Júpiter. La inserción en la órbita de Júpiter se llevará a cabo en octubre de 2029, y será seguida por un recorrido enorme dentro del sistema de Júpiter. En este viaje se visitarán las lunas Europa y Calisto, así como el propio Júpiter. El punto final del viaje será la luna Ganímedes, la luna más grande de todo el Sistema Solar, donde JUICE entrará en órbita a su alrededor, en septiembre de 2032. El fin de la misión está previsto para Junio de 2033, cuando la nave se quede sin combustible.

El objetivo de JUICE es estudiar las condiciones que pueden haber conducido a la aparición de ambientes habitables entre las lunas heladas de Júpiter, con especial énfasis en Ganímedes, Europa y Calisto, que probablemente albergan océanos de agua líquida bajo sus superficies. La luna Ganímedes es, de hecho, el principal objetivo de la misión. Es la luna más grande de todo el Sistema Solar, y creemos que hay quizás 6 veces más agua líquida dentro de esta luna que en la propia Tierra. Además, Ganímedes es extremadamente fascinante ya que también tiene un campo magnético interno, como la Tierra o Mercurio, que hace que este cuerpo sea realmente único.

Todas estas lunas están en muy alta medida influenciadas por Júpiter. El campo magnético de Júpiter es enorme, así como su atmósfera y su propia radiación, lo que hace que tenga un efecto enorme en sus lunas. La misión JUICE también se centrará en caracterizar la diversidad de procesos que tienen lugar en el sistema de Júpiter y que son los responsables de cambios en la superficie



Fig.8. Imagen artística de la ESA sobre la misión JUICE en Júpiter y de sus lunas Ío, Europa, Calisto, y Ganímedes.

y atmósferas de Ganímedes, Europa y Calisto, que van desde escalas cortas de tiempo, hasta escalas geológicas. Estudiar el sistema de Júpiter a nivel global es esencial para comprender el origen y la evolución del sistema, y sobre todo, para saber si alguna de sus lunas pudiera tener un ambiente (en su interior) favorable para la vida.

Más allá de Júpiter: Saturno, los gigantes de hielo y los confines del Sistema Solar.

Más allá de Júpiter nos encontramos con otro de los planetas gigantes y gaseosos de nuestro Sistema Solar: Saturno. Durante más de 13 años (2004-2017), la sonda Cassini nos descubrió la belleza y grandeza de este planeta, así como de sus anillos y lunas. La misión de Cassini a Saturno, un esfuerzo conjunto de la NASA, la ESA y la agencia espacial italiana, ha sido uno de los esfuerzos más ambiciosos en la exploración del espacio planetario jamás montado. Entre los grandes descubrimientos de Cassini cabe destacar una estimación más precisa de la duración del día en Saturno (¡no es fácil saber cuánto dura un día en un planeta que es todo gas!), el estudio detallado de sus campos gravitatorios y magnéticos, el minucioso análisis de la evolución de la corriente de chorro de forma hexagonal (mejor conocida como "el hexágono") que ocupa todo el polo norte del planeta con un vórtice giratorio en su centro, o el descubrimiento de decenas de nuevas lunas, algunas de ellas entre sus propios anillos.

El sistema de Saturno, como lo es el de Júpiter, es realmente fascinante. Pero si hay algo que de verdad cambió el ritmo de la exploración espacial, fue el estudio de dos de sus lunas: Titán y Encelado. La misión Cassini también llevó consigo una sonda europea llamada Huygens, que se lanzó en paracaídas a la superficie de Titán en 2005. Titán es la luna más grande de Saturno y la segunda luna más grande del Sistema Solar. Gracias a Cassini sabemos que es un mundo oceánico, con mares y con una atmósfera de 1.5 bares de metano, lo que hace que esta luna sea muy parecida a la Tierra. De hecho, podríamos decir que Titán es como una Tierra metida en una nevera. La sonda Huygens aterrizó en su superficie en enero de 2005 enviándonos resultados espectaculares. Hasta la fecha, Huygens es la nave que más lejos de la Tierra ha aterrizado en otro mundo. Sin embargo, fue Encelado la luna que nos dio la gran sorpresa. La sonda Cassini descubrió que esta pequeña luna alberga también un océano bajo su superficie, y que su agua líquida es expulsada al espacio a través de grietas en forma de gigantescos géiseres. Este es el último cuerpo del Sistema Solar donde podríamos pensar en detectar agua líquida. Por eso su descubrimiento fue simplemente asombroso, poniendo en el punto de mira a las lunas de hielo de los planetas gigantes como mundos más habitables que nunca. La ciencia revelada por Cassini nos ayudará durante décadas en la búsqueda de vida también en otros Sistemas Solares, ya que las lecciones aprendidas de esta misión se pueden fácilmente extrapolar a los planetas que orbitan otras estrellas, los llamados “exoplanetas”.

Más allá de Saturno, nos encontramos con los gigantes de hielo Urano y Neptuno. Ambos son muy interesantes pero, por desgracia, poco conocidos. La mayor parte de la información que tenemos sobre ellos proviene de los sobrevuelos de la sonda



Fig.9. Imagen artística de Neptuno visto desde Tritón.

Imagen obtenida desde:

https://www.flickr.com/photos/a_responsibility_to_awe/11945380963

sobrevuelos de la sonda Voyager-2 en 1986 y 1989, y de escasas observaciones de telescopios. Urano y Neptuno representan una clase muy especial de planetas: menos masivos que Júpiter y Saturno, pero con una composición que hace pensar que es de rocas, hielo, y gas. Urano tiene la particularidad de tener un eje de rotación alineado con su eje orbital, es decir, gira tumbado, lo que es algo muy inusual. Estudios recientes sugieren que

esto se deba quizás a una gran colisión que tuvo lugar en el pasado lejano. Sus sistemas lunares también son muy interesantes, en particular la luna Tritón alrededor de Neptuno, que se cree que es un objeto capturado del Cinturón de Kuiper (un cinturón de asteroides más allá de la órbita de Plutón). Tritón se cree que pueda ser otro mundo acuático con agua líquida en su interior.

Realmente necesitamos una misión para explorar estos mundos, lo que sin duda, nos traerá grandes descubrimientos tal y como hizo Cassini la pasada década.

Aún más allá de los planetas gigantes de hielo nos encontramos con Plutón y sus lunas, exploradas por la nave espacial de la NASA New Horizons durante un memorable sobrevuelo en julio de 2015. A pesar de considerarlo como un planeta enano en los confines del Sistema Solar, Plutón ha resultado ser un mundo mucho más complejo e interesante de lo previsto. La sonda New Horizons nos descubrió que este planeta enano alberga el mayor glaciar del Sistema Solar, multitud de cadenas montañosas de varios kilómetros de altura, y una atmosfera tenue pero muy fría (unos -230°C) que alberga multitud de partículas orgánicas.

Finalmente, nos gustaría hablar del último y más lejano objeto explorado por el ser humano hasta la fecha, Ultima Thule. Este cuerpo, a priori un clásico del cinturón de Kuiper, ha sido recientemente visitado por la sonda New Horizons en enero de 2019. Las últimas fotos recibidas muestran un cuerpo formado de dos piezas, como si fuera un muñeco de nieve, parece que hubieran colisionado hace tiempo. Aún sabemos muy poco sobre este objeto ya que el equipo de New Horizons se encuentra actualmente analizando todos los datos e imágenes recibidas del sobrevuelo, pero “stay tuned! (¡estad atentos!)” porque en los próximos meses nuestra curiosidad ¡tendrá respuesta!



Fig.10. Imágenes de Plutón (arriba) y de Ultima Thule (abajo) obtenidas por la sonda de la NASA New Horizons. Créditos: NASA/JHUAPL/SwRI

Exoplanetas, nuestro siguiente destino.

Los planetas que orbitan alrededor de otras estrellas se conocen como exoplanetas. Son muy difíciles de detectar con telescopios porque el resplandor brillante de las estrellas que orbitan les oculta. Para detectarlos, es necesario el uso de diversas técnicas, como por ejemplo, a través del tránsito del planeta delante de su estrella. En esos casos, cuando el planeta transita frente a la estrella, bloquea un poco la luz de la estrella, que a partir del cambio de brillo se puede obtener el tamaño del planeta. El primer exoplaneta se descubrió en 1995, y desde entonces, el desarrollo de su exploración ha sido espectacular.

Gracias a misiones como Kepler, actualmente sabemos que los exoplanetas son muy comunes en el universo. Sin embargo, la parte más importante de esta nueva forma de exploración está aún por llegar: la evidencia

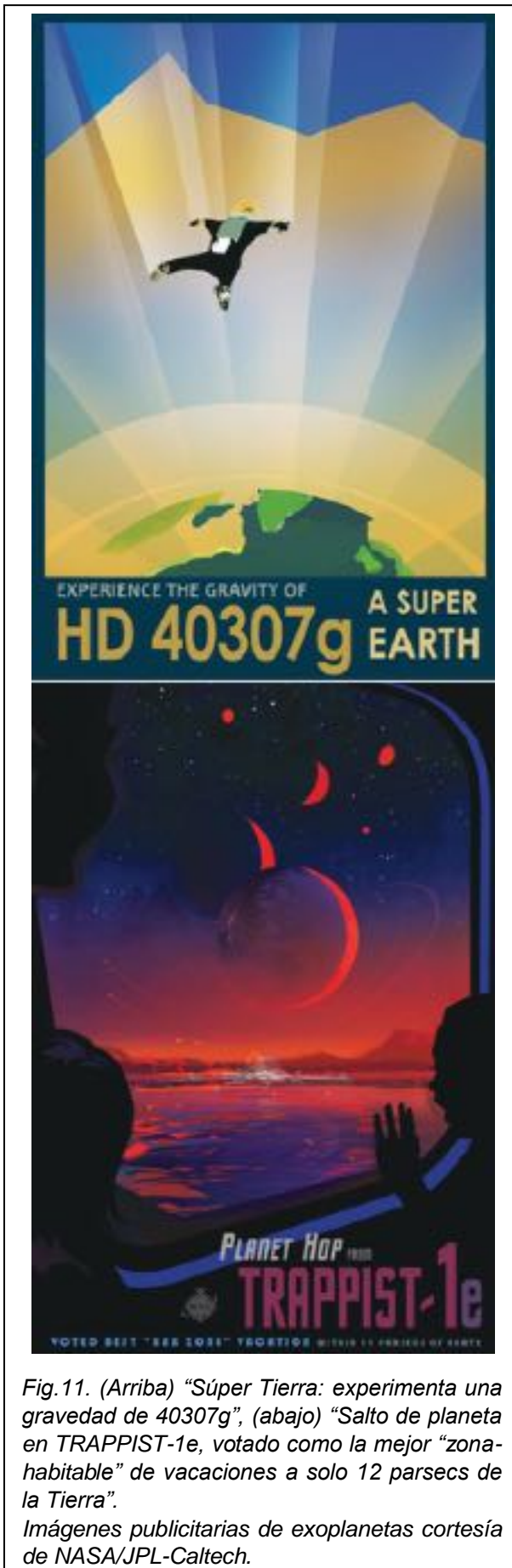


Fig.11. (Arriba) “Súper Tierra: experimenta una gravedad de 40307g”, (abajo) “Salto de planeta en TRAPPIST-1e, votado como la mejor “zona-habitable” de vacaciones a solo 12 parsecs de la Tierra”.

Imágenes publicitarias de exoplanetas cortesía de NASA/JPL-Caltech.

de un mundo hospitalario para la vida. La clave de esta búsqueda está en encontrar la zona habitable de cada estrella. Tanto las estrellas como los planetas son de muy diversos tipos y tamaños, y por tanto la interacción de estos factores determina el alcance y la influencia de la zona habitable. Si tomamos como ejemplo la Tierra, podemos definir muy claramente donde está la zona habitable de todas las estrellas con características similares a nuestro Sol que viene marcado por la distancia y la temperatura adecuadas para que exista agua líquida.

La exploración de exoplanetas es un campo muy joven pero con una gran comunidad científica detrás que crece cada día y que ve en ellos al futuro de la exploración espacial. En solo 20 años de exploración de exoplanetas, ya hemos encontrado más de 5000 planetas, algunos de ellos pueden considerarse súper-Tierras, es decir, planetas que poseen entre una y diez veces la masa de la Tierra. También, hemos encontrado sistemas planetarios como el de la estrella TRAPPIST-1, que está compuesto de siete planetas terrestres con tamaño similar a la Tierra y de los cuales, tres orbitan dentro de su zona habitable. Todo parece indicar que en las próximas décadas podremos encontrar vida en algún exoplaneta, y para ellos se preparan todas las agencias espaciales del planeta. Sólo en Europa, la ESA planea el lanzamiento de cuatro misiones durante la próxima década. Estas son: CHEOPS (CHaracterising ExOPlanets Satellite) en 2019, James Webb Space Telescope en 2021, la misión PLATO (PLANetary

Transits and Oscillations of stars) en 2026, y la misión ARIEL (Atmospheric Remote-sensing Infrared Exoplanet Large-survey) en 2028.

Conclusiones.

La exploración espacial esta más de moda que nunca. Las próximas décadas nos traerán grandes descubrimientos dentro y fuera de nuestro Sistema Solar, y quién sabe si incluso la detección de vida en otros mundos. Como se dice en inglés “stay tuned! (¡estad atentos!)” porque el futuro ¡promete más que nunca!

*¡Nunca te olvides,
que nosotros exploramos el Sistema Solar
a bordo de la maravillosa nave espacial
llamada Tierra!*

Beatriz Sanchez-Cano

Doctora en Ciencias Físicas.

Post-Doctoral Research Associate, University of Leicester, Reino Unido.

Olivier Witasse

Doctor en Ciencias Físicas.

Científico de Proyecto de la misión JUper ICy moons Explorer (JUICE) de la Agencia Espacial Europea.

Agencia Espacial Europea, Noordwijk, Países Bajos.

Este libro se terminó de editar en Abril de 2019, siendo un proyecto de difusión gratuito.
Agradecemos inmensamente la aportación de los autores. Todos ellos han colaborado desinteresadamente.



INSPIRACIÓN DE PORTADA

Cielo de Marte: Ilustración sobre fotografía tomada por el Rover Curiosity el 31 de Enero de 2014. [\[Archivo NASA\]](#)