

# Moda, fe y fantasía en la nueva física del universo

Roger Penrose

## Índice

Moda, fe y fantasía en la nueva física del universo

Prefacio

¿Son la moda, la fe o la fantasía relevantes para la ciencia fundamental?

1. Moda

1.1. La elegancia matemática como fuerza motriz

1.2. Algunas modas físicas del pasado

1.3. Los antecedentes de la teoría de cuerdas en la física de partículas

1.4. El principio de superposición en la QFT

1.5. El poder de los diagramas de Feynman

1.6. Las ideas clave originales de la teoría de cuerdas

1.7. El tiempo en la relatividad general de Einstein

1.8. La teoría de gauge del electromagnetismo de Weyl

1.9. Libertad funcional en modelos de Kaluza-Klein y de cuerdas

1.10. ¿Obstáculos cuánticos a la libertad funcional?

1.11. Inestabilidad clásica de la teoría de cuerdas supradimensional

1.12. El estatus de moda de la teoría de cuerdas

1.13. Teoría M

1.14. Supersimetría

1.15. AdS/CFT

1.16. Mundos de branas y el paisaje

2. Fe

2.1. La revelación cuántica

- 2.2. Max Planck y  $E = hn$
- 2.3. La paradoja onda-partícula
- 2.4. Niveles cuántico y clásico: **C, U y R.**
- 2.5. Función de onda de una partícula puntual
- 2.6. Función de onda de un fotón
- 2.7. Linealidad cuántica
- 2.8. La medición cuántica
- 2.9. La geometría del espín cuántico
- 2.10. Entrelazamiento cuántico y efectos EPR
- 2.11. Libertad funcional cuántica
- 2.12. Realidad cuántica
- 2.13. Reducción objetiva del estado cuántico: ¿un límite para la fe cuántica?
- 3. Fantasía
  - 3.1. El Big Bang y las cosmologías FLRW
  - 3.2. Agujeros negros e irregularidades locales
  - 3.3. La segunda ley de la termodinámica
  - 3.4. La paradoja del Big Bang
  - 3.5. Horizontes, volúmenes comóviles y diagramas conformes
  - 3.6. La fenomenal precisión del Big Bang
  - 3.7. ¿Entropía cosmológica?
  - 3.8. Energía del vacío
  - 3.9. Cosmología inflacionaria
  - 3.10. El principio antrópico
  - 3.11. Otras cosmologías fantásticas
- 4. ¿Una nueva física para el universo?
  - 4.1. Teoría de twistores: ¿una alternativa a las cuerdas?
  - 4.2. ¿Do van los cimientos cuánticos?
  - 4.3. ¿Cosmología chiflada conforme?
  - 4.4. Una coda personal
- Apéndice matemático
  - A.1. Exponentes reiterados

- A.2. Libertad funcional de los campos
- A.3. Espacios vectoriales
- A.4. Bases vectoriales, coordenadas y duales
- A.5. Matemáticas de las variedades
- A.6. Variedades en física
- A.7. Fibrados
- A.8. Libertad funcional mediante fibrados
- A.9. Números complejos
- A.10. Geometría compleja
- A.11. Análisis armónico
- Bibliografía
- Agradecimientos
- Créditos de las ilustraciones
- Índice alfabético
- Notas
- Sobre este libro
- Sobre Roger Penrose
- Créditos

## **Prefacio**

### **¿SON LA MODA, LA FE O LA FANTASÍA RELEVANTES PARA LA CIENCIA FUNDAMENTAL?**

Este libro surge a partir de la recopilación de tres conferencias que ofrecí en la Universidad de Princeton en octubre de 2003 invitado por Princeton University Press. Es muy posible que me precipitase al proponer a la editorial el título para esas tres charlas —Moda, fe y fantasía en la nueva física del universo—, que es también el título de este libro, pero expresaba genuinamente cierta desazón que sentía entonces en relación con determinadas tendencias que formaban parte del pensamiento de la época sobre las leyes físicas que rigen el universo en el que vivimos. Ha pasado más de una década, pero esas cuestiones, y mucho de lo que dije sobre ellas, parecen ser, en su mayor parte, al menos tan relevantes hoy como en su día. Debo confesar que di esas conferencias con cierta aprensión, pues intentaba expresar unos puntos de vista que podrían no resultar muy acordes con los de muchos de los distinguidos expertos presentes allí.

Cada uno de los nombres que dan título a este libro, «moda», «fe» y «fantasía», remite a una cualidad que podría parecer antagónica con los procedimientos que se suelen considerar apropiados para la búsqueda de los principios que subyacen al comportamiento del universo a sus niveles más básicos. De hecho, idealmente, sería muy razonable afirmar que influencias como las de la moda, la fe o la fantasía deberían estar por completo ausentes de la actitud mental de quienes dedican todos sus esfuerzos a la búsqueda de las bases fundamentales de nuestro universo. A fin de cuentas, no cabe duda de que la propia naturaleza no tiene mucho interés

en los caprichos efímeros de las modas humanas; ni tampoco la ciencia debería entenderse como una fe, pues sus dogmas están sometidos a un escrutinio continuo y sujetos a los rigores del examen experimental, y se abandonan en el mismo momento en que surge un conflicto convincente con la realidad de la naturaleza tal y como la descubrimos. A su vez, la fantasía es sin duda el territorio de ciertas zonas de la ficción y del entretenimiento, donde no se considera esencial que se preste demasiada consideración a los requisitos de coherencia con la observación, a la lógica estricta o siquiera al mero sentido común. De hecho, si se puede demostrar que una teoría científica propuesta está demasiado influida por los dictados de las modas, por el seguimiento incondicional de una fe sin base experimental o por las tentaciones románticas de la fantasía, entonces es nuestro deber poner de manifiesto dichas influencias y mantener alejado de estas a quien, quizá sin ser consciente de ello, pudiera estar expuesto a ellas.

No obstante, no pretendo mostrarme por completo negativo respecto a estas cualidades, pues se puede argumentar que cada uno de estos términos contiene algo genuinamente valioso. Al fin y al cabo, es poco probable que una teoría de moda posea tal estatus solo por motivos sociológicos. Deben existir sin duda muchas cualidades positivas para que infinidad de investigadores se congreguen en torno a un área de estudio muy a la moda, y es poco probable que sea el mero deseo de formar parte de una multitud lo que mantiene a esos investigadores tan fascinados por lo que probablemente sea un campo de estudio en extremo difícil, dificultad que a menudo radica en la naturaleza altamente competitiva de las actividades de moda.

Cabe señalar aquí una cuestión más en relación con la investigación en física teórica que pueda estar de moda pero que dista de ser una descripción posible del mundo (de hecho, como veremos, a menudo está en flagrante contradicción con las observaciones actuales). Quienes trabajan en esos ámbitos podrían haber encontrado en ellos una enorme gratificación si los hechos observados se hubiesen revelado en mayor sintonía con sus propias visiones del mundo, pero con frecuencia parecen relativamente imperturbables cuando los hechos son menos complacientes de lo que les habría gustado. Esto no es algo descabellado pues, en buena medida, estas investigaciones son meramente exploratorias, y se entiende que estos trabajos pueden permitir ganar experiencia y que esto resultará a la larga en beneficio para el descubrimiento de mejores teorías que concuerden mejor con el funcionamiento real del universo que conocemos.

Cuando de lo que se trata es de la fe extrema en ciertos dogmas científicos que suelen expresar algunos investigadores, es probable que exista una poderosa razón que la justifique, aun cuando la fe está en la aplicabilidad de dicho dogma en circunstancias alejadas de las situaciones originales donde erigieron sus cimientos gracias a un potente respaldo observacional. Podemos tener confianza en que las espléndidas teorías físicas del pasado seguirán proporcionando una enorme precisión incluso cuando, en determinadas circunstancias, han sido reemplazadas por teorías mejores que extienden su precisión o su ámbito de aplicabilidad. Así sucedió sin duda cuando la magnífica teoría de la gravitación de Newton fue sustituida por la de Einstein, o cuando la hermosa teoría electromagnética de la luz de Maxwell fue desplazada por su versión cuantizada, que

permitía comprender la faceta de la luz como partículas (fotones). En cada caso, la teoría anterior conservaría su fiabilidad, siempre que sus limitaciones se tuviesen debidamente en cuenta.

¿Y la fantasía? Esta debería ser todo lo contrario de aquello a lo que aspira la ciencia. No obstante, veremos que existen ciertos aspectos de la naturaleza del universo real que son tan excepcionalmente extraños (aunque no siempre se reconozca del todo que lo son) que si no nos permitimos lo que podría interpretarse como escandalosas fantasías no tendremos ninguna posibilidad de llegar a aceptar lo que bien podría ser una realidad en apariencia extraordinariamente fantástica.

En los tres primeros capítulos, ilustraré estas tres cualidades que dan título al libro con tres teorías —o familias de teorías— muy conocidas. No he escogido campos de la física de importancia relativamente menor, sino lo que son de hecho peces gordos en el océano de la actividad actual en física teórica. En el capítulo 1, he decidido tratar algo que está aún muy de moda: la teoría de cuerdas (o teoría de supercuerdas, o sus generalizaciones, como la teoría M, o el aspecto más de moda de toda esta línea general de trabajo, como es el orden de cosas conocido como correspondencia AdS/CFT). La fe que abordaré en el capítulo 2 es un pez aun más gordo: el dogma según el cual los procedimientos de la mecánica cuántica deben seguirse servilmente, con independencia de lo grandes o masivos que sean los elementos físicos sobre los que se apliquen. Y, en cierto sentido, el tema del capítulo 3 es el pez más gordo de todos, ya que en él nos interesaremos por el mismísimo origen del universo que conocemos y haremos un repaso de algunas de las propuestas de aparente pura fantasía que se han

formulado para abordar determinadas peculiaridades ciertamente inquietantes que se han puesto de manifiesto gracias a observaciones bien fundamentadas de los primerísimos instantes del universo en su conjunto.

Por último, en el capítulo 4 expongo varias ideas personales, con la intención de poner de manifiesto que existen caminos alternativos que podrían muy bien transitarse. Veremos, no obstante, que seguir las sendas que sugiero no estaría exento de cierta ironía. Está, qué duda cabe, la ironía de seguir la vía que yo mismo prefiero para entender la física fundamental, una vía que le presentaré brevemente al lector en §4.1. Es un camino jalonado por la teoría de twistores, en cuyo desarrollo he tenido una participación fundamental y a la que la comunidad física apenas había prestado atención durante cerca de cuarenta años, pero que, como veremos, ahora empieza a adquirir cierta notoriedad en relación con la teoría de cuerdas.

En cuanto a la fe suprema e incuestionable en la mecánica cuántica que parece profesar la mayor parte de la comunidad física, esta se ha visto reafirmada por experimentos notables como los de Serge Haroche y David Wineland, que recibieron un merecido reconocimiento al ser galardonados con el Premio Nobel de Física en 2012. Asimismo, el Nobel de 2013 para Peter Higgs y François Englert, por su contribución a la predicción de lo que se ha dado en conocer como el bosón de Higgs, es una asombrosa confirmación no solo de las ideas particulares que (junto con otros cuantos más, en particular Tom Kibble, Gerald Guralnik, Carl R. Hagen y Robert Brout) habían propuesto sobre el origen de las masas de las partículas, sino también de muchos de los aspectos fundamentales de la propia teoría cuántica (de campos). Pero, como señalo en §4.2, todos los experimentos de

elevada sofisticación que se han llevado a cabo hasta ahora distan considerablemente del nivel de desplazamiento de masa (tal y como se propone en §2.13) que sería necesario para que se pueda prever con visos de realidad el cuestionamiento de nuestra fe cuántica. Sin embargo, hoy en día se están preparando otros experimentos pensados para detectar dicho nivel de desplazamiento de masa, que en mi opinión podrían contribuir a resolver algunos de los conflictos profundos que existen entre la mecánica cuántica actual y determinados principios físicos también aceptados, como los de la relatividad general de Einstein. En §4.2 señalo un grave conflicto entre la mecánica cuántica actual y el principio, fundamental para Einstein, de equivalencia entre campos gravitatorios y aceleraciones. Quizá los resultados de esos experimentos debiliten la inquebrantable fe mecanocuántica que tan extendida parece estar. Por otra parte, cabe preguntarse por qué debería alguien depositar más fe en el principio de equivalencia de Einstein que en los procedimientos fundamentales de la mecánica cuántica, que han sido sometidos a muchísimas más pruebas. Esta es desde luego una buena pregunta. Y se podría argumentar sin duda que es necesaria tanta fe para aceptar el principio de Einstein como los propios de la mecánica cuántica. Se trata de un asunto que bien podría zanjarse experimentalmente (como ironía final) en un futuro no muy lejano.

Por lo que se refiere a los niveles de fantasía que ha alcanzado la cosmología actual, en §4.3 sugiero que existe un sistema denominado cosmología cíclica conforme (CCC), que yo mismo propuse en 2005, que es, en ciertos aspectos, aún más fantástico que las extraordinarias propuestas que veremos en el capítulo 3, algunas de las cuales han pasado a formar parte de casi todos los debates contemporáneos

sobre los primerísimos instantes del universo. Pero los análisis observacionales actuales parecen comenzar a poner de manifiesto que la CCC tiene cierta base en los hechos físicos reales. Cabe sin duda confiar en que pronto se disponga de evidencia experimental inequívoca susceptible de convertir lo que podría ser mera fantasía, de un tipo u otro, en una imagen convincente de la naturaleza factual del universo real. De hecho, puede señalarse que, a diferencia de las modas de la teoría de cuerdas o de la mayoría de los sistemas teóricos que buscan minar nuestra fe absoluta en los principios de la mecánica cuántica, esas fantásticas propuestas que se plantean para describir el origen mismo del universo ya se han sometido a detallados exámenes observacionales, como el de la exhaustiva información que proporcionan los satélites COBE, WMAP y Planck, o el de los resultados de las observaciones BICEP2 en el Polo Sur publicados en marzo de 2014. Mientras escribo esto, existen importantes dificultades para la interpretación de estos últimos, pero no debería tardarse mucho en resolverlas. Quizá tengamos pronto pruebas mucho más claras que permitan escoger definitivamente entre varias de estas teorías fantásticas rivales o alguna otra teoría que aún está por ser planteada.

Al tratar de abordar estos asuntos de manera satisfactoria (pero no demasiado técnica), he debido hacer frente a un obstáculo muy concreto y fundamental: el problema de las matemáticas y su papel central en cualquier teoría física que pretenda describir seriamente la naturaleza con una mínima profundidad. Los argumentos críticos que expondré en este libro, que buscan establecer que las modas, la fe y la fantasía están en efecto ejerciendo una influencia indebida sobre el progreso de la ciencia fundamental, deben basarse, en una

medida significativa, en verdaderas objeciones técnicas y no en meras preferencias emocionales, lo cual nos obligará a utilizar cierta cantidad de matemáticas. Pero este texto no pretende ser un discurso técnico, accesible solo para expertos en matemáticas o física, ya que la intención inequívoca es que puedan leerlo con provecho personas legas. Así pues, trataré de limitar el contenido técnico a un mínimo razonable. No obstante, hay una serie de conceptos matemáticos que resultarán muy útiles para la plena comprensión de varios asuntos esenciales que pretendo tratar. Por ese motivo, he incluido once apartados matemáticos bastante básicos en un apéndice, cuyo contenido no es muy técnico pero que podrían, si fuera necesario, ayudar a los no expertos a alcanzar una comprensión más profunda de muchos de los asuntos principales.

Los dos primeros de estos apartados (§§A.1 y A.2) solo contienen ideas muy sencillas, aunque relativamente poco conocidas, y no emplean notación difícil. No obstante, desempeñan un papel especial en relación con muchas de las argumentaciones que aparecen en el libro, en particular en lo tocante a las propuestas de moda que se discuten en el capítulo 1. Cualquier lector que desee comprender el asunto crítico que allí se analiza debería, en algún momento, ahondar en el material de §§A.1 y A.2, pues contiene la clave de mi argumento contra la idea de que existen de verdad dimensiones adicionales en nuestro universo. Esta supradimensionalidad es una de las afirmaciones fundamentales de casi toda la teoría de cuerdas moderna y sus variantes principales. Mis argumentos críticos se centran en la creencia actual, debida a la teoría de cuerdas, de que el número de dimensiones del espacio físico debe ser superior

a las tres que experimentamos directamente. El asunto clave que planteo aquí es el de la libertad funcional, y en §A.8 esbozo un argumento algo más completo para clarificar la idea básica. El concepto matemático en cuestión hunde sus raíces en la obra del gran matemático francés Élie Cartan, que data de principios del siglo XX, pero no goza de mucho aprecio entre los físicos teóricos de hoy en día a pesar de su gran relevancia en relación con la verosimilitud de las ideas físicas sobre dimensiones adicionales.

La teoría de cuerdas y sus variantes modernas han progresado en muchos aspectos en los años transcurridos desde estas conferencias en Princeton, y ha aumentado considerablemente su grado de precisión técnica. En modo alguno pretendo arrogarme ninguna erudición sobre dichos avances, aunque he revisado buena parte de ese material. Lo que más me interesa no son esos detalles sino esclarecer si ese trabajo nos permite de verdad avanzar hacia la comprensión del mundo físico real en el que vivimos. Muy en particular, veo pocos intentos (por no decir ninguno) de abordar la cuestión de la excesiva libertad funcional que surge de la supradimensionalidad espacial que se presupone. De hecho, ninguno de los trabajos sobre teoría de cuerdas que haya leído menciona siquiera este problema. Esto me resulta en cierta medida sorprendente, y no solo porque esta cuestión fuese esencial en una de mis conferencias en Princeton, hace ya más de una década. Ya lo había destacado en una charla que di en el congreso que se celebró en la Universidad de Cambridge en conmemoración del sexagésimo cumpleaños de Stephen Hawking, en enero de 2002, ante un público entre el que se encontraban varios eminentes teóricos de cuerdas, y de la que posteriormente se escribieron reseñas.

Debo señalar aquí algo importante. A menudo, los físicos cuánticos rechazan el problema de la libertad funcional al afirmar que solo tiene validez en la física clásica, y las dificultades que presenta para las teorías supradimensionales suelen desdeñarse sumariamente con un argumento que busca demostrar la irrelevancia de estos aspectos en situaciones mecanocuánticas. En §1.10 expongo mi razonamiento principal contra este argumento básico, y animo encarecidamente a que lo lean los defensores de la supradimensionalidad. Confío en que, al repetir dichos argumentos aquí y al desarrollarlos en determinados contextos físicos (§§1.10, 1.11, 2.11 y A.11), consiga que se tengan en cuenta en obras futuras.

Los demás apartados del apéndice presentan brevemente los espacios vectoriales, las variedades, el análisis armónico, los números complejos y su geometría. Estos temas serán con toda seguridad conocidos para los expertos, pero es posible que a quienes los desconozcan este material introductorio y mesurado les resulte útil para comprender por completo los aspectos más técnicos de este libro. En todas mis descripciones, he optado por no incluir una introducción en profundidad a las ideas del cálculo diferencial (o integral), pues opino que, aun cuando entender correctamente el cálculo sería beneficioso para los lectores, quienes carezcan de antemano de este conocimiento ganarían poco con un apresurado apartado sobre el tema. No obstante lo cual, en §A.11 he considerado apropiado incluir un somero repaso de los operadores y las ecuaciones diferenciales, para así poder explicar algunos asuntos de relevancia, en varios sentidos, para el hilo argumental que recorre el libro.

# 1

## Moda

### 1.1. LA ELEGANCIA MATEMÁTICA COMO FUERZA MOTRIZ

Como he mencionado en el prefacio, los temas que se tratan en este libro los desarrollé a partir de tres conferencias que di, invitado por Princeton University Press, en la Universidad de Princeton en octubre de 2003. Al dirigirme a un público tan experto como la comunidad científica de Princeton, mi nerviosismo alcanzó quizá el punto álgido cuando llegó el momento de hablar de las modas, porque el campo que había decidido utilizar como ejemplo, la teoría de cuerdas y algunas de sus derivaciones, se había desarrollado hasta su máxima expresión en Princeton probablemente más que en cualquier otro lugar del mundo. Además, se trata de un asunto muy técnico, y no puedo pretender tener competencia sobre muchos de sus importantes aspectos, pues mi familiaridad con los detalles técnicos es limitada, sobre todo habida cuenta de mi condición de persona ajena a este campo. Me pareció, no obstante, que no debía permitir que esta carencia me amedrentase, pues si se considerase que solo los expertos están en condiciones de hacer comentarios críticos sobre el asunto, es probable que las críticas se limitaran a detalles relativamente técnicos, y ciertos aspectos más amplios resultarían, sin duda, en buena medida ignorados.

Desde que pronuncié estas conferencias, se han publicado tres análisis muy críticos de la teoría de cuerdas: *Not Even Wrong*, de Peter Woit; *Las dudas de la física en el siglo XXI*, de Lee Smolin, y *Farewell to Reality. How Fairytale Physics Betrays the Search for Scientific Truth*, de Jim Baggott. Desde

luego, tanto Woit como Smolin han tenido más experiencia directa que yo de la comunidad de los teóricos de cuerdas y su condición de vanguardia de la moda. También se han publicado en este tiempo (antes que las tres obras citadas) mis críticas a la teoría de cuerdas, en el capítulo 31 y partes del 34 de *El camino a la realidad*, pero mis comentarios críticos eran quizá ligeramente más favorables a reconocer un papel físico a la teoría de cuerdas que esas otras obras. La mayoría de mis comentarios serán de naturaleza general, y apenas relacionados con cuestiones muy técnicas.

Permítanme decir en primer lugar algo general (y quizá evidente). Observamos que el asombroso progreso que la teoría física ha experimentado a lo largo de varios siglos ha dependido de sistemas matemáticos extremadamente precisos y sofisticados. Es evidente, por lo tanto, que cualquier avance significativo debe de nuevo depender de algún armazón matemático particular. Para que cualquier nueva teoría física que se proponga pueda mejorar lo logrado hasta el momento y haga predicciones precisas e inequívocas que superen lo que había sido posible con anterioridad, también debe estar basada en un sistema matemático bien definido. Además, cabe pensar, para que se trate de una teoría matemática aceptable debería tener sentido matemático (lo cual significa que, en efecto, debería ser matemáticamente coherente). A partir de un sistema incoherente uno podría, en principio, deducir cualquier resultado que desease.

Pero la coherencia interna es en realidad un criterio muy fuerte, y resulta que no muchas propuestas de teorías físicas —ni siquiera entre las de mayor éxito en el pasado— son en realidad internamente coherentes. A menudo, hay que apelar a sólidas consideraciones físicas para que la teoría pueda

aplicarse apropiadamente de manera inequívoca. Los experimentos son también, desde luego, primordiales para una teoría física, y someter una teoría a una prueba experimental es algo muy distinto de buscar su coherencia lógica. Ambas son importantes, pero en la práctica a menudo se observa que a los físicos no les preocupa tanto alcanzar la plena coherencia matemática interna de una teoría si esta parece encajar con los hechos físicos. Así ha sucedido, en buena medida, con la mecánica cuántica, como veremos en el capítulo 2 (y en §1.3). La primerísima obra sobre este tema, la trascendental propuesta de Max Planck para explicar el espectro de frecuencias de la radiación electromagnética en equilibrio con la materia a una temperatura fija (el espectro del cuerpo negro; véanse §§2.2 y 2.11), tuvo que emplear una representación híbrida que no era realmente coherente [Pais, 2005]. Tampoco puede decirse que la vieja teoría cuántica del átomo, tan magníficamente presentada por Niels Bohr en 1913, fuese un sistema coherente. Los avances posteriores de la mecánica cuántica han permitido erigir un edificio matemático de gran sofisticación, en el que el deseo de coherencia matemática ha sido una poderosa fuerza motriz. Pero en la teoría actual aún persisten problemas de coherencia que esta no aborda debidamente, como veremos más adelante, sobre todo en §2.13. Sin embargo, es el respaldo experimental, sobre una amplia variedad de fenómenos físicos, lo que constituye el fundamento pesado de la teoría cuántica. Los físicos no suelen preocuparse demasiado por detalles de incoherencia matemática u ontológica si la teoría, cuando se aplica con el debido criterio y mediante cálculos minuciosos, sigue proporcionando respuestas que concuerdan plenamente con los resultados de la observación —a menudo con una precisión

extraordinaria— obtenidos a través de experimentos delicados y precisos.

La situación de la teoría de cuerdas es por completo distinta. En este caso, no parece que exista ningún resultado experimental que la respalde. Se suele argumentar que esto no es algo sorprendente, ya que la teoría de cuerdas tal y como está formulada en la actualidad, en gran medida como una teoría de la gravedad cuántica, se preocupa fundamentalmente por la llamada escala de Planck de distancias diminutas (o, al menos, por valores próximos a dichas distancias), unas 10-15 o 10-16 veces más pequeñas (10-16 significa, por supuesto, un factor de una diezmilésima de una millonésima de una millonésima parte), y por lo tanto por energías unas 10<sup>15</sup> o 10<sup>16</sup> veces mayores que las alcanzables en los experimentos actuales. (Debe señalarse que, según los principios básicos de la relatividad, una distancia pequeña equivale en esencia a un tiempo pequeño, mediante la velocidad de la luz, y, según los principios de la mecánica cuántica, un tiempo pequeño es básicamente equivalente a una energía grande, mediante la constante de Planck; véanse §§2.2 y 2.11.) Uno debe sin duda afrontar el hecho de que, por potentes que sean nuestros aceleradores de partículas actuales, las energías que se prevé que alcancen son muchísimo menores que las que parecen tener relevancia directa en teorías como la de cuerdas moderna que tratan de aplicar los principios de la mecánica cuántica a los fenómenos gravitatorios. Pero esta situación difícilmente puede considerarse satisfactoria para una teoría física, ya que el soporte experimental es el criterio definitivo para que esta perviva o sea desechada.

Desde luego, podría ser que estemos entrando en una nueva fase de la investigación básica en física fundamental, en la

que los requisitos de coherencia matemática sean primordiales y en que, cuando dichos requisitos (junto con una coherencia con los principios establecidos previamente) resulten insuficientes, haya que recurrir a criterios adicionales de elegancia matemática. Aunque podría parecer acientífico apelar a esos ideales estéticos en una búsqueda completamente objetiva de los principios físicos que rigen el funcionamiento del universo, es llamativo lo fructíferos —esenciales, de hecho— que tales juicios estéticos han resultado ser en muchas ocasiones. En física existen muchos ejemplos en los que bellas ideas matemáticas han resultado ser la base de avances fundamentales en nuestro conocimiento. El gran físico teórico Paul Dirac [1963] fue muy claro sobre la importancia del criterio estético en su descubrimiento de la ecuación del electrón, y también en su predicción de las antipartículas. Sin duda, la ecuación de Dirac ha resultado ser absolutamente fundamental para la física básica, y su atractivo estético es muy apreciado. Lo mismo sucede con la idea de las antipartículas, derivadas del análisis profundo que el propio Dirac hizo de su ecuación del electrón.

Sin embargo, es muy difícil valorar de manera objetiva el papel del criterio estético. Es habitual que algún físico considere que un determinado sistema es muy bello mientras que otro no está en absoluto de acuerdo. En el mundo de la física teórica, igual que sucede en el arte o el diseño de ropa, las modas pueden tener una influencia desproporcionada en lo que respecta a criterios estéticos.

Debería quedar claro que la cuestión del criterio estético en física es más sutil que lo que se conoce como la navaja de Occam, la eliminación de complicaciones innecesarias. De hecho, la decisión de cuál de entre dos teorías enfrentadas

es en efecto «la más sencilla», y quizá por lo tanto la más elegante, no tiene por qué ser una cuestión evidente. Por ejemplo, ¿es la relatividad general de Einstein una teoría sencilla o no? ¿Es más sencilla o más complicada que la teoría de la gravedad de Newton? ¿O es la teoría de Einstein más sencilla o más complicada que la propuesta en 1894 por Asaph Hall (veintiún años antes de que Einstein presentase su teoría de la relatividad general), que es como la de Newton pero en la que la ley de la inversa del cuadrado se sustituye por otra en la que la fuerza gravitatoria entre una masa  $M$  y otra masa  $m$  es  $GmM/r^{2,00000016}$ , en lugar de  $GmM/r^2$  como en la de Newton? La teoría de Hall se proponía explicar la ligera desviación observada en el desplazamiento del perihelio del planeta Mercurio respecto a las predicciones de la teoría de Newton, algo conocido desde aproximadamente 1843. (El perihelio es el punto más cercano al Sol que un planeta alcanza mientras recorre su órbita [Roseveare, 1982].) Esta teoría también se ajustaba al movimiento de Venus ligeramente mejor que la de Newton. En cierto sentido, la teoría de Hall es solo ligeramente más complicada que la de Newton, aunque depende de cuánta complicación adicional considera uno que supone reemplazar el número «2», fácil y sencillo, por «2,00000016». No cabe duda de que esta sustitución implica una pérdida de elegancia matemática, pero, como ya se ha dicho, estos juicios contienen un elemento importante de subjetividad. Puede que sea más relevante el hecho de que existen ciertas propiedades matemáticas elegantes que se deducen de la ley de la inversa del cuadrado (que expresan, básicamente, la conservación de las «líneas de flujo» de la fuerza gravitatoria, algo que no sería exactamente cierto en la teoría de Hall). Pero, de nuevo, esta podría considerarse una cuestión estética cuyo significado físico no debería exagerarse.

¿Qué hay de la relatividad general de Einstein? Sin duda, a la hora de examinar en detalle las consecuencias de la teoría, su aplicación a sistemas físicos concretos resulta mucho más difícil que aplicar la teoría de Newton (o incluso la de Hall). Cuando se escriben de forma explícita, las ecuaciones son mucho más complicadas en la teoría de Einstein, e incluso es difícil hacerlo con todo detalle. Además, son enormemente más difíciles de resolver, y en la teoría de Einstein existen muchas no linealidades que no aparecen en la de Newton (que tienden a invalidar los sencillos argumentos de conservación del flujo que ya tuvo que abandonar la teoría de Hall). (Véanse §§A.4 y A.11 para el significado de linealidad, y §2.4 para su función especial en la mecánica cuántica.) Aun más grave es el hecho de que la interpretación física de la teoría de Einstein depende de la eliminación de efectos espurios que surgen al escoger unas determinadas coordenadas, a pesar de que dicha elección no debería tener ninguna relevancia en esta teoría. En términos prácticos, no cabe duda de que la teoría de Einstein suele ser mucho más difícil de manejar que la teoría gravitatoria de Newton (o incluso que la de Hall).

No obstante, en un sentido importante la teoría de Einstein es de hecho muy sencilla, quizá incluso más (o más «natural») que la de Newton: depende de la teoría matemática de la geometría de Riemann (o, más precisamente, pseudoriemanniana, como veremos en §1.7), de 4-variedades de curvatura arbitraria (véase también §A.5). No es fácil dominar este conjunto de técnicas matemáticas, pues necesitamos entender qué es un tensor y cuál es el propósito de dichas magnitudes, y, en particular, cómo construir el objeto tensorial  $R$ , el tensor de curvatura de Riemann, a partir del tensor métrico,  $g$ , que define la

geometría. A continuación, mediante una contracción y una inversión de la traza, podremos construir el tensor de Einstein,  $G$ . Sin embargo, las ideas geométricas generales en que se basa el formalismo son razonablemente fáciles de entender, y, una vez que se comprenden los ingredientes de este tipo de geometría curva, resulta que solo existe una familia muy reducida de ecuaciones posibles (o verosímiles) que sean compatibles con los requisitos generales propuestos, tanto físicos como geométricos. Entre estas posibilidades, la más sencilla de todas nos da la famosa ecuación del campo de Einstein de la relatividad general,  $G = 8\pi\gamma T$  (donde  $T$  es el tensor de masa-energía de la materia y  $\gamma$  es la constante gravitatoria de Newton —según la propia definición de este último, de manera que ni siquiera el  $8\pi$  es en realidad una complicación sino simplemente cuestión de cómo prefiramos definir  $\gamma$ ).

Existe una pequeña —y muy sencilla— modificación de la ecuación del campo de Einstein que puede hacerse dejando intactos los requisitos esenciales del sistema, como es la inclusión de un número constante  $L$ , la denominada constante cosmológica (que Einstein introdujo en 1917 por motivos que más tarde desechó), de manera que las ecuaciones de Einstein, con  $L$ , se convierten ahora en  $G = 8\pi\gamma T + Lg$ . Hoy en día, la magnitud  $L$  suele denominarse energía oscura, presumiblemente para considerar la posibilidad de generalizar la teoría de Einstein de forma que  $L$  pueda variar. Existen, no obstante, fuertes restricciones matemáticas que dificultan tales consideraciones, y en §§3.1, 3.7, 3.8 y 4.3, donde  $L$  desempeñará un papel importante para nosotros, me limitaré a analizar situaciones en las que  $L$  no varía. La constante cosmológica tendrá una relevancia considerable en el capítulo 3 (y también en §1.15). De hecho,

observaciones relativamente recientes apuntan a todas luces a la presencia física real de  $L$  con un minúsculo valor positivo (en apariencia constante). La evidencia de que  $L > 0$  —o, posiblemente, de una forma más general de energía oscura— es ahora muy impresionante y ha ido aumentando desde las observaciones iniciales de Perlmutter et al. [1999], Riess et al. [1998] y sus colaboradores, que condujeron en 2011 a la concesión del Premio Nobel de Física a Saul Perlmutter, Brian P. Schmidt y Adam G. Riess. Este  $L > 0$  tiene relevancia inmediata solo en las escalas cosmológicas muy remotas, y las observaciones de movimientos celestiales a una escala más local se pueden tratar adecuadamente según la ecuación original de Einstein,  $G = 8\pi\gamma T$ , más sencilla. Se ha comprobado que esta ecuación proporciona una precisión inaudita a la hora de modelar el comportamiento de los cuerpos celestes bajo la influencia de la gravedad; el valor observado de  $L$  no tiene un efecto apreciable sobre tales dinámicas locales.

A este respecto, es de la máxima importancia histórica el sistema estelar binario PSR1913+16, uno de cuyos componentes es un púlsar que emite señales electromagnéticas con una frecuencia muy precisa que se reciben en la Tierra. El movimiento de cada estrella alrededor de la otra, al tratarse muy nítidamente de un efecto puro gravitatorio, puede modelarse mediante la relatividad general con una precisión extraordinaria que puede argumentarse que, en su conjunto, es de alrededor de  $10^{14}$ , acumulada a lo largo de unos cuarenta años. Este periodo de cuarenta años equivale aproximadamente a 109 segundos, por lo que una precisión de uno sobre  $10^{14}$  implica una concordancia entre observación y teoría de hasta alrededor de  $10^{-5}$  (una cienmilésima) de segundo en dicho periodo, que es,

asombrosamente, justo el resultado que se observa. Más recientemente, otros sistemas [Kramer et al., 2006] de uno o incluso dos púlsares tienen el potencial de incrementar bastante esta precisión cuando dichos sistemas han sido observados durante un periodo de tiempo comparable al que se lleva observando PSR1913+16.

Pero afirmar que esta cifra de 10<sup>-14</sup> es una medida de la precisión experimental de la relatividad general es algo que suscita cierto debate. De hecho, las masas y los parámetros orbitales concretos deben calcularse a partir de los movimientos observados, en lugar de ser cifras procedentes de la teoría o de observaciones independientes. Además, mucha de esta precisión extraordinaria ya existe en la teoría gravitatoria de Newton.

Pero aquí nos interesan las teorías gravitatorias en su conjunto, y la de Einstein incorpora las deducciones de la de Newton (que dan como resultado las órbitas elípticas de Kepler, etc.) como primera aproximación, aunque introduce varias correcciones a las órbitas keplerianas (incluido el desplazamiento del perihelio), y por último una pérdida de energía del sistema que concuerda exactamente con una notable predicción de la relatividad general: que un sistema tan masivo en movimiento acelerado debería perder energía a través de la emisión de ondas gravitatorias, ondulaciones en el espaciotiempo que constituyen el análogo gravitacional de las ondas electromagnéticas (esto es, de la luz) que los cuerpos cargados eléctricamente emiten cuando experimentan un movimiento acelerado. Constituye una asombrosa confirmación adicional de la existencia y la forma precisa de esta radiación gravitatoria el anuncio [Abbott et al., 2016] de su observación directa en el detector de ondas gravitatorias LIGO, lo que también proporciona una excelente

evidencia directa de otra de las predicciones de la relatividad general: la existencia de agujeros negros, a la que volveré en §3.2 y que también trataré en apartados posteriores del capítulo 3 y en §4.3.

Debe recalarse que esta precisión es muy superior —en un factor adicional de alrededor de 10<sup>8</sup> (es decir, de cien millones) o más— a la observacionalmente alcanzable cuando Einstein formuló su teoría gravitatoria. Se podría afirmar que la precisión observada de la teoría gravitatoria de Newton es del orden de uno sobre 10<sup>7</sup>. Por lo tanto, la precisión de «uno sobre 10<sup>14</sup>» de la relatividad general ya estaba «ahí», en la naturaleza, antes de que Einstein formulase su teoría. Pero esa precisión adicional (de un factor de alrededor de cien millones), al desconocerla Einstein, no pudo haber tenido ninguna influencia en su formulación de la teoría. Así, este nuevo modelo matemático de la naturaleza no era una construcción artificial inventada meramente en un intento de encontrar la mejor teoría que encajase con los datos; el esquema matemático ya existía, a todas luces, en los propios entresijos de la naturaleza. Esta simplicidad —o elegancia— matemática, o como queramos describirla, forma parte verdaderamente del comportamiento de la naturaleza; no se trata tan solo de que nuestras mentes estén predispuestas a quedar impresionadas ante tal belleza matemática.

Por otra parte, cuando tratamos deliberadamente de utilizar el criterio de la belleza matemática al formular nuestras teorías, nos dejamos engañar con facilidad. La relatividad general es sin duda una teoría muy hermosa, pero ¿cómo juzga uno la elegancia de las teorías físicas en general? Personas diferentes poseen criterios estéticos muy distintos. No resulta necesariamente evidente que el punto de vista de

una sobre lo que es elegante coincida con el de las demás, o que el criterio estético de una persona sea superior o inferior al de otra, a la hora de formular una buena teoría física. A menudo, además, la belleza inherente a una teoría no es evidente en un primer momento y puede revelarse tiempo después, cuando avances técnicos posteriores ponen de manifiesto la profundidad de su estructura matemática. La dinámica de Newton es un buen ejemplo. Buena parte de la indudable belleza del esquema de Newton se manifestó mucho más adelante, gracias a los extraordinarios trabajos de grandes matemáticos como Euler, Lagrange, Laplace y Hamilton (de lo que dan cuenta expresiones como ecuaciones de Euler-Lagrange, operador laplaciano, lagrangianos y hamiltonianos, que son ingredientes esenciales de la teoría física moderna). La tercera ley de Newton, por ejemplo, que afirma que a toda acción se opone una reacción igual y de dirección opuesta, ocupa un lugar central en la formulación lagrangiana de la física moderna. No me sorprendería descubrir que la belleza que tan a menudo se afirma que existe en las buenas teorías modernas es en muchas ocasiones y en cierta medida algo post hoc. El propio éxito de una teoría física, tanto experimental como matemático, puede contribuir significativamente a las cualidades estéticas que se le descubren a posteriori. De todo esto cabe deducir que valorar propuestas de teorías físicas a partir de sus supuestas cualidades estéticas es seguramente problemático o ambiguo. Es sin duda más seguro formarse la opinión sobre una nueva teoría en función de su concordancia con las observaciones actuales y de su capacidad predictiva.

Aun así, en lo que se refiere al soporte experimental, con frecuencia los experimentos cruciales no son factibles,

como sucede con las altas energías por completo prohibitivas que, según se dice, deberían alcanzar las diferentes partículas —extravagantemente superiores a las que logran los aceleradores de partículas actuales (véase §1.10)— para someter como es debido a la prueba del experimento cualquier teoría de la gravedad cuántica. Otras propuestas experimentales más modestas también pueden resultar igualmente irrealizables, debido quizá al coste de los experimentos o a su dificultad intrínseca. Incluso en el caso de los experimentos de mayor éxito, con mucha frecuencia sucede que los experimentadores recopilan cantidades ingentes de datos y el problema es de muy distinta índole, pues se trata entonces de extraer de ese cenagal de datos algún pedazo clave de información. Esto es desde luego así en la física de partículas, en que potentes aceleradores y colisionadores de partículas producen hoy día masas de información, y empieza a serlo también en cosmología, donde las observaciones modernas de la radiación de fondo de microondas (CMB, por sus siglas en inglés) generan enormes cantidades de datos (véanse §§3.4, 3.9 y 4.3). Muchos de estos no se consideran demasiado informativos, ya que simplemente confirman lo que ya se sabía como resultado de experimentos anteriores. Se hace necesario un monumental esfuerzo de procesamiento estadístico para extraer minúsculos residuos —la información novedosa que buscan los experimentalistas— que podrían confirmar o refutar alguna propuesta teórica.

Debería señalarse aquí que, con toda probabilidad, este procesamiento estadístico es muy específico de la teoría en cuestión y está dirigido a detectar cualquier ligero efecto adicional que esta pudiera predecir. Es muy probable que un conjunto de ideas radicalmente distintas, que se alejen de

manera sustancial de las modas actuales, permanezcan sin ser sometidas a prueba aun cuando una respuesta definitiva bien podría estar oculta en los datos ya existentes, y no salga a la luz porque los procedimientos estadísticos que los físicos han adoptado están demasiado orientados hacia la teoría actual. Veremos lo que resulta ser un sorprendente ejemplo de esto en §4.3. Aun cuando está claro que podría extraerse estadísticamente información definitiva de un marasmo de datos fiables, el exorbitante tiempo de computación que esto podría requerir constituye en ocasiones una enorme barrera para la realización efectiva del análisis, en particular cuando este compite directamente con otros proyectos más en boga.

Aún más relevante es el hecho de que los propios experimentos suelen ser enormemente caros, y es probable que su diseño específico esté dirigido a comprobar teorías que no se apartan de las ideas convencionales. Cualquier sistema teórico que se aleje de forma demasiado radical del consenso general tendrá dificultades para reunir los fondos suficientes para someterse a las debidas pruebas. A fin de cuentas, un aparato experimental muy caro requiere que muchos comités de expertos establecidos aprueben su construcción, y es probable que estos expertos participasen en el desarrollo de las perspectivas actuales.

En relación con este asunto podemos considerar el Gran Colisionador de Hadrones (LHC, por sus siglas en inglés) de Ginebra (Suiza), cuya construcción se completó en 2008. Está formado por un túnel de 27 kilómetros bajo dos países (Francia y Suiza) que empezó a operar en 2010. Ahora se le atribuye el descubrimiento de la hasta entonces esquiva partícula de Higgs, de gran importancia en la física de partículas, sobre todo en relación con su papel a la hora de

asignar masa a las partículas que experimentan la interacción débil. El Premio Nobel de Física de 2013 se otorgó a Peter Higgs y François Englert por su contribución al innovador trabajo que predijo la existencia y las propiedades de esta partícula.

Se trata de un logro indiscutiblemente espectacular, y no tengo ninguna intención de menospreciar su indudable importancia. Sin embargo, el LHC es un ejemplo muy apropiado. La manera en que se analizan los encuentros entre partículas a muy altas energías requiere la presencia de detectores extraordinariamente caros, pensados para recabar información en relación con la teoría predominante en la física de partículas. Puede que no resulte nada fácil obtener información relevante para otras ideas no convencionales sobre la naturaleza básica de las partículas fundamentales y sus interacciones. En general, las propuestas que se alejan de manera drástica de una perspectiva dominante pueden tener muchas más dificultades para obtener una financiación adecuada, e incluso para ser sometidas a prueba mediante experimentos definitivos.

Otro factor importante es que los estudiantes de doctorado, cuando buscan un problema en el que trabajar para obtener el título, suelen estar sujetos a restricciones estrictas en cuanto a los temas de investigación que se consideran adecuados. Los estudiantes investigadores que trabajan en los campos que están menos de moda, aun cuando sus investigaciones culminen satisfactoriamente en tesis doctorales, pueden toparse con grandes dificultades para obtener después puestos académicos, por mucho que sea su talento, conocimiento u originalidad. Los puestos de trabajo son limitados y la financiación para la investigación,

difícil de conseguir. La mayoría de las veces, a los directores de investigación les interesa sobre todo desarrollar ideas en cuya promoción ellos mismos han intervenido, y lo más probable es que correspondan a campos que ya estén de moda. Además, un director interesado en desarrollar una idea ajena a la corriente dominante puede ser reacio a animar a un potencial estudiante a trabajar en ese ámbito, por la desventaja que podría suponer para este a la hora de medirse posteriormente en un mercado laboral muy competitivo en el que quienes posean experiencia en campos más en boga tendrán una clara ventaja.

Los mismos problemas surgen en lo tocante a la financiación de los proyectos de investigación. Las propuestas en campos en boga tienen una probabilidad mucho más alta de recibir la aprobación (véase también §1.12). De nuevo, las propuestas serán valoradas por expertos reconocidos, que con toda probabilidad trabajan en campos que ya están de moda, y a los que ellos mismos habrán contribuido de forma significativa. Los proyectos que se desvíen mucho de las normas actualmente aceptadas, aunque estén bien pensados y sean muy originales, es harto probable que no reciban apoyo. Además, no es solo cuestión de lo limitada que sea la financiación disponible, ya que la influencia de las modas parece ser especialmente relevante en Estados Unidos, donde la disponibilidad de fondos para la investigación científica continúa siendo relativamente elevada.

Debemos mencionar, por supuesto, que la mayoría de las líneas de investigación menos en boga tendrán una probabilidad considerablemente menor de llegar a ser teorías de éxito que cualquiera de las que ya están de moda. Una perspectiva radicalmente novedosa tendrá, en la

inmensa mayoría de los casos, pocas posibilidades de dar lugar a una propuesta viable. Ni que decir tiene que, como en el caso de la relatividad general de Einstein, todas estas perspectivas radicales deben de antemano concordar con lo que ya está establecido experimentalmente y, de no ser así, puede que no se necesite una cara prueba experimental para rechazar las ideas inadecuadas. Pero, para las propuestas teóricas que concuerdan con todos los experimentos previos, y para las que no existen en la actualidad perspectivas de confirmación o refutación experimental —quizá por motivos como los que se acaban de exponer—, parece que deberíamos volver a la coherencia matemática, la aplicabilidad general y criterios estéticos a la hora de juzgar la verosimilitud y relevancia de una propuesta de teoría física. Es en estas circunstancias cuando el papel de las modas puede alcanzar proporciones excesivas, por lo que debemos tener mucho cuidado de no permitir que el hecho de que una determinada teoría esté más o menos en boga nos obnubile sobre su verosimilitud física.

## **1.2. ALGUNAS MODAS FÍSICAS DEL PASADO**

Esto es particularmente importante para las teorías que pretenden sondear los fundamentos mismos de la realidad física, como la moderna teoría de cuerdas, y debemos cuidarnos mucho de atribuir demasiada verosimilitud a tales teorías como consecuencia de su popularidad. Sin embargo, antes de tratar ideas físicas actuales, resultará instructivo mencionar algunas de las teorías científicas que fueron populares en el pasado pero que hoy en día no nos tomamos en serio. Son muy numerosas y estoy convencido de que la mayoría de los lectores sabrán muy poco de la mayor parte de ellas, por el simple motivo de que, si no nos las tomamos en serio, es muy poco probable que las estudiemos (a menos,

claro está, que seamos buenos historiadores de la ciencia, lo que no es el caso de la mayoría de los físicos). Permítanme al menos mencionar unas cuantas de las más conocidas.

En particular, está la antigua teoría griega que relacionaba los sólidos platónicos con lo que entonces se consideraban los elementos básicos de la sustancia material, tal y como se representa en la figura 1-1. En ella, el fuego se representa mediante el tetraedro regular; el aire, como el octaedro; el agua, a través del icosaedro, y la tierra, como el cubo. Además, posteriormente se añadió el éter celeste (también llamado firmamento o quintaesencia), del que se suponía que estaban compuestos los cuerpos celestes, y para cuya representación se utilizó el dodecaedro regular. Parece que fueron los antiguos griegos quienes formularon este tipo de visión —o al menos muchos de ellos la suscribían—, y supongo que se podría considerar una teoría de moda en aquella época.

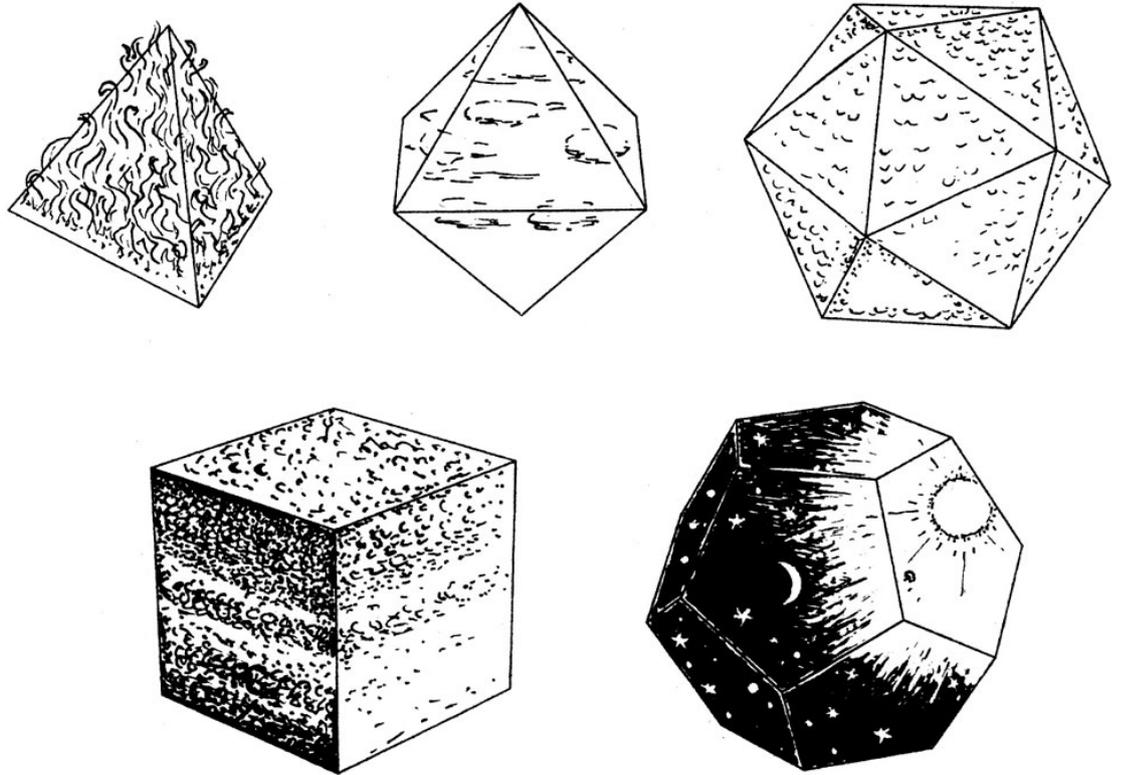


FIGURA 1-1. Los cinco elementos de la antigua Grecia: fuego (tetraedro), aire (octaedro), agua (icosaedro), tierra (cubo) y éter (dodecaedro).

Al principio, solo tenían los cuatro elementos (fuego, aire, agua y tierra), y esta colección de entidades primitivas parecía encajar con las cuatro formas poliédricas perfectamente regulares que se conocían por aquel entonces, pero cuando, más tarde, se descubrió el dodecaedro, la teoría hubo de ampliarse para acomodar este nuevo poliedro. En consecuencia, la sustancia celeste que componía cuerpos tan supuestamente perfectos como el Sol, la Luna y los planetas, y las esferas de cristal a las que estaban supuestamente fijados, se incorporó al sistema de los poliedros (para los griegos, esta sustancia se regía por

leyes muy distintas de las existentes en la Tierra y poseía un movimiento aparentemente eterno, en lugar de seguir la tendencia universal entre las sustancias habituales a ralentizarse y detenerse). Quizá de esto pueda extraerse alguna lección sobre la manera en que incluso las sofisticadas teorías modernas, tras haberse presentado en un principio en una forma supuestamente definitiva, pueden después sufrir alteraciones importantes y sus doctrinas originales, ser estiradas hasta extremos impensados, a la vista de nuevas evidencias teóricas o experimentales. Tal y como yo lo veo, la idea que tenían los antiguos griegos era que, de alguna manera, las leyes que regían el movimiento de las estrellas, los planetas, la Luna y el Sol eran muy diferentes de las que afectaban a las cosas en la Tierra. Tuvieron que llegar Galileo, con su visión de la relatividad del movimiento, y más tarde Newton, con su teoría de la gravitación universal —muy influida por las ideas de Kepler sobre las órbitas de los planetas—, para que se tomase conciencia de que los cuerpos celestes se rigen por las mismas leyes que los terrestres.

Cuando entré en contacto por primera vez con las ideas de los antiguos griegos, me parecieron pura fantasía romántica, sin ninguna base matemática (y menos aún física). Pero más tarde supe que tras ellas había algo más de teoría de lo que yo había imaginado en un principio. Algunas de estas formas poliédricas pueden dividirse en partes y, a continuación, recombinarse debidamente para crear otras (como, por ejemplo, dos cubos se pueden cortar para crear dos tetraedros y un octaedro). Esto podría estar relacionado con el comportamiento físico de esos diferentes elementos y usarse como un modelo geométrico de las transiciones que pueden producirse entre ellos. Al menos, aquí tenemos una

atrevida e imaginativa suposición sobre la naturaleza de la sustancia material, que no era descabellada en una época en que aún se sabía tan poco sobre la verdadera naturaleza y el comportamiento del material físico. Este fue uno de los primeros intentos de encontrar una base para los materiales reales en términos de una estructura matemática elegante — muy en consonancia con lo que los físicos teóricos aún aspiran a hacer hoy en día— en la que las consecuencias teóricas del modelo pudiesen ser puestas a prueba frente al comportamiento físico real. También intervenían claramente criterios estéticos, y sin duda estas ideas parecen haber sido del interés de Platón. Pero ni que decir tiene que los detalles no soportaron bien el paso del tiempo, pues de lo contrario no habríamos desechado tan atractiva propuesta matemática.

Consideremos unas cuantas cuestiones más. El modelo ptolemaico del movimiento de los planetas —en el que la Tierra estaba fija y situada en el centro del cosmos— tuvo un éxito extraordinario y fue indiscutible durante muchos siglos. Los movimientos del Sol, la Luna y los planetas debían entenderse en términos de epiciclos, que permitían explicar los movimientos planetarios mediante la superposición de varios movimientos circulares uniformes. Aunque el sistema debía ser bastante complicado para proporcionar un buen encaje con las observaciones, no carecía por completo de elegancia matemática, y ofrecía una teoría razonablemente predictiva del movimiento futuro de los planetas. Debe señalarse que, cuando se consideran los movimientos externos respecto a una Tierra estacionaria, los epiciclos carecen de una verdadera justificación. Los movimientos que observamos directamente desde el punto de vista terrestre implican la composición de la rotación de la Tierra (lo que da

lugar a un movimiento circular aparente del firmamento alrededor del eje polar terrestre), junto con los movimientos aparentes generales del Sol, la Luna y los planetas, que están aproximadamente limitados al plano elíptico, y que nosotros vemos muy aproximadamente como un movimiento circular respecto a un eje distinto. Por sólidos motivos geométricos, ya percibimos algo sobre la naturaleza general de los epiciclos —movimientos circulares sobre otros movimientos circulares—, por lo que no era descabellado suponer que esta idea podría extenderse de manera más general a los movimientos más detallados de los planetas.

Además, los propios epiciclos poseen una geometría interesante, y el mismo Ptolomeo era un excelente geómetra. En su trabajo astronómico, empleó un elegante y potente teorema geométrico que posiblemente descubriera él mismo, pues ahora lleva su nombre. (Dicho teorema afirma que la condición para que cuatro puntos —A, B, C y D— en un plano estén situados en un círculo —tomados en ese orden cíclico— es que las distancias entre ellos satisfagan que  $AB \cdot CD + BC \cdot DA = AC \cdot BD$ .) Esta fue la teoría aceptada del movimiento de los planetas durante alrededor de catorce siglos, hasta que fue sustituida, y al final desmontada por completo, gracias al maravilloso trabajo de Copérnico, Galileo, Kepler y Newton, y ahora se la considera completamente errónea. No obstante, debe describirse sin duda como una teoría de moda, y tuvo un éxito extraordinario durante unos catorce siglos (desde mediados del siglo II hasta mediados del XVI) al ser capaz de dar cuenta muy aproximadamente de todas las observaciones de movimientos planetarios (con la introducción cada cierto tiempo de las debidas mejoras), hasta que, a finales del siglo XVI, Tycho Brahe obtuvo mediciones más precisas.

Otra teoría famosa en la que ahora no creemos, aunque estuvo muy en boga durante más de un siglo, entre 1667 (cuando Joshua Becher la propuso) y 1778 (fecha en que fue refutada por Antoine Lavoisier), fue la del flogisto, según la cual cualquier sustancia inflamable contenía un elemento denominado flogisto, que era emitido a la atmósfera durante el proceso de combustión. La teoría del flogisto explicaba la mayoría de los datos relativos a la combustión que se conocían por aquel entonces, tales como el hecho de que, cuando esta tenía lugar en un recipiente sellado razonablemente pequeño, solía finalizar antes de que se hubiese consumido todo el material combustible, lo cual se achacaba a que el aire dentro del contenedor se saturaba de flogisto y era incapaz de absorber más. Curiosamente, Lavoisier fue el responsable de otra teoría de moda pero falsa, según la cual el calor era una sustancia material a la que bautizó como calórico. Esta teoría fue desacreditada en 1798 por el conde de Rumford (sir Benjamin Thompson).

En cada uno de estos dos ejemplos principales, el éxito de una teoría puede entenderse por su estrecha relación con la teoría más satisfactoria que la sustituyó. En el caso de la dinámica ptolemaica, podemos movernos a la más satisfactoria visión heliocéntrica de Copérnico mediante una sencilla transformación geométrica consistente en tomar el Sol como centro de los movimientos, en lugar de la Tierra. Al principio, cuando todo se describía en función de los epiciclos, este cambio apenas supuso ninguna diferencia, salvo por el hecho de que la visión heliocéntrica parecía mucho más sistemática, pues en ella los movimientos más rápidos correspondían a los planetas más cercanos al Sol [Gingerich, 2004; Sobel, 2011], y en ese momento existía una equivalencia básica entre ambos sistemas. Pero, cuando

Kepler descubrió sus tres leyes del movimiento planetario elíptico, la situación cambió por completo, ya que una descripción geocéntrica de este tipo de movimiento no tenía sentido desde un punto de vista geométrico. Las leyes de Kepler fueron la clave que despejó el camino hacia la visión newtoniana, extraordinariamente precisa y general, de la gravedad universal. No obstante, la visión geocéntrica podría no parecer tan extravagante hoy en día como en el siglo XIX, habida cuenta del principio general de covariancia de la relatividad general de Einstein (véanse §§1.7, A.5 y 2.13), que nos permite considerar descripciones mediante sistemas de coordenadas muy incómodos (como uno geocéntrico en el que las coordenadas de la Tierra no varían con el paso del tiempo) como igualmente válidas. De la misma manera, se podría establecer una estrecha correspondencia entre la teoría del flogisto y la visión moderna de la combustión, según la cual el hecho de que un material arde normalmente implica que se consuma oxígeno de la atmósfera, y en la que el flogisto podría verse simplemente como «oxígeno negativo». Esto nos proporciona una traducción razonablemente coherente entre la visión del flogisto y la que ahora se considera convencional. Pero cuando las precisas mediciones de Lavoisier demostraron que el flogisto debería poseer una masa negativa, esa representación comenzó a perder apoyo. Sin embargo, el de «oxígeno negativo» no es un concepto tan absurdo en el contexto de la física de partículas moderna, en que se supone que en la naturaleza cada tipo de partícula (incluidas las compuestas) posee una antipartícula; así pues, un «antiátomo de oxígeno» encaja a la perfección dentro de la teoría moderna. Aunque, eso sí, ¡tendría una masa negativa!

A veces, teorías que han estado pasadas de moda durante un tiempo pueden volver a tomarse en consideración como consecuencia de acontecimientos posteriores. Un buen ejemplo es una idea que lord Kelvin (William Thompson) propuso alrededor de 1867, según la cual se podía entender que los átomos (las partículas elementales de su época) estaban compuestos por diminutas estructuras similares a nudos. La idea concitó por aquel entonces una atención considerable, y basándose en ella el matemático J. G. Tait comenzó el estudio sistemático de los nudos. Pero la teoría no permitió establecer ninguna correspondencia clara con el comportamiento físico real de los átomos, por lo que cayó en el olvido. Sin embargo, más recientemente ideas de este tipo han vuelto a encontrar respaldo, en parte gracias a su relación con los conceptos de la teoría de cuerdas. La teoría matemática de los nudos también ha experimentado un resurgimiento desde aproximadamente 1984 a raíz del trabajo de Vaughan Jones, cuyas ideas fundamentales tenían su origen en reflexiones teóricas en el marco de la teoría cuántica de campos [Jones, 1985; Skyrme, 1961]. Con posterioridad, Edward Witten [1989] empleó los métodos de la teoría de cuerdas para obtener un tipo de teoría cuántica de campos (denominada teoría topológica cuántica de campos) que, en cierto sentido, enmarca estos nuevos desarrollos dentro de la teoría matemática de nudos.

Como ejemplo de resurgimiento de una idea mucho más antigua sobre la naturaleza del universo a gran escala, podría mencionar —aunque no del todo en serio— una curiosa coincidencia que se produjo aproximadamente cuando impartí la conferencia en Princeton en la que se basa este capítulo en particular (el 17 de octubre de 2003). En esa charla, hice referencia a la antigua idea griega según la cual

el éter estaba relacionado con el dodecaedro regular. Por aquel mismo entonces, y sin que yo lo supiera, se publicaron en los periódicos informaciones sobre una propuesta, obra de Luminet et al. [2003], según la cual la geometría espacial tridimensional del cosmos podría en realidad poseer una topología algo complicada, surgida a partir de la identificación (con un giro) de las caras opuestas de un dodecaedro regular (sólido). De forma que, en cierto sentido, la idea platónica de un cosmos dodecaédrico estaba reviviendo en nuestra época.

La ambiciosa idea de una teoría del todo, que englobaría todos los procesos físicos, incluida una descripción de todas las partículas de la naturaleza y sus interacciones físicas, se ha discutido a menudo en los últimos años, en particular en relación con la teoría de cuerdas. Se trataría de tener una teoría completa del comportamiento físico, basada en una cierta idea de partículas elementales y/o campos, que actuarían según ciertas fuerzas u otros principios dinámicos que regirían con precisión los movimientos de todos los elementos constituyentes. Esto también podría interpretarse como el resurgimiento de una idea antigua, como veremos más adelante.

Al mismo tiempo que Einstein le daba la forma definitiva a su teoría general de la relatividad, hacia finales de 1915, el matemático David Hilbert propuso su propio método para deducir las ecuaciones de campo de la teoría de Einstein,[1] utilizando lo que se conoce como principio variacional. (Este tipo de procedimiento, muy general, hace uso de las ecuaciones de Euler-Lagrange, obtenidas a partir de un lagrangiano, un concepto potente al que se hace referencia explícita en §1.1; véase, por ejemplo, Penrose [2004: cap. 20]; en adelante, me referiré a este libro como ECalR.) Einstein,

en su enfoque más directo, formuló explícitamente sus ecuaciones de manera que ponían de manifiesto cómo se comportaría el campo gravitatorio (descrito en función de la curvatura del espaciotiempo) bajo la influencia de su «fuente», esto es, las densidades totales de masa/energía de todas las partículas, todos los campos de materia, etc., recogidas en forma del tensor de energía  $T$  (que aparece en §1.1).

Einstein no ofreció ninguna prescripción específica sobre las ecuaciones concretas que debían regir el comportamiento de estos campos de materia, pues se suponía que estas debían tomarse de alguna otra teoría específica para los campos de materia concretos en cuestión. En particular, uno de estos campos sería el electromagnético, cuya descripción vendría dada por las maravillosas ecuaciones que el gran físico matemático escocés James Clerk Maxwell descubrió en 1864 y que unificaron los campos eléctrico y magnético, explicando así la naturaleza de la luz y de buena parte de las fuerzas que gobiernan la constitución interna de los materiales ordinarios. En este contexto, esto se consideraría materia, y desempeñaría su debido papel en  $T$ . Además, también podrían intervenir otros tipos de campo, y toda clase de otras especies de partículas, que se regirían por las ecuaciones correspondientes, fueran cuales fuesen, y también se considerarían materia y contribuirían a  $T$ . Los detalles de todo esto no eran importantes para la teoría de Einstein, y no se especificaron.

Por su parte, en su propuesta Hilbert buscaba abarcar más, y ahora podríamos considerarla una teoría del todo. El campo gravitatorio se describiría exactamente de la misma manera que en la propuesta de Einstein, pero en lugar de dejar sin especificar el término fuente  $T$ , como había hecho Einstein,

Hilbert propuso que este término fuente debería ser el de una teoría muy particular que estaba en boga en la época, conocida como teoría de Mie [Mie, 1908, 1912a,b, 1913]. Esto implicaba una modificación no lineal de la teoría electromagnética de Maxwell, y había sido propuesto por Gustav Mie como una vía para incorporar todos los aspectos de la materia. En consecuencia, se suponía que la propuesta global era una teoría completa tanto de la materia (incluido el electromagnetismo) como de la gravedad. Por aquel entonces, aún no se comprendían las fuerzas fuerte y débil de la física de partículas, pero la propuesta de Hilbert podría en efecto haber sido entendida como lo que ahora solemos denominar una teoría del todo. Sin embargo, no creo probable que muchos físicos actuales hayan siquiera oído hablar de la teoría de Mie, que tan de moda estuvo en otra época, y menos aún del hecho de que formó parte explícitamente de la versión de Hilbert de la relatividad general como una teoría del todo. Esa teoría no desempeña ningún papel en la comprensión moderna de la materia. Quizá de esto pueda extraerse una lección de prudencia para los teóricos de hoy en día, decididos a proponer sus propias teorías del todo.

### **1.3. LOS ANTECEDENTES DE LA TEORÍA DE CUERDAS EN LA FÍSICA DE PARTÍCULAS**

Una de estas propuestas teóricas es la teoría de cuerdas, y muchos físicos teóricos actuales aún consideran que ofrece un camino claro hacia una teoría del todo. La teoría de cuerdas tuvo su origen en una serie de ideas que, cuando las oí por primera vez alrededor de 1970 (en boca de Leonard Susskind), me parecieron sorprendentemente atractivas y de una naturaleza particularmente convincente. Pero, antes de describirlas, tendría que situarlas en su debido contexto.

Deberíamos tratar de entender por qué sustituir la noción de una partícula puntual por un pequeño lazo o curva en el espacio, como de hecho proponía la versión original de la teoría de cuerdas, podría parecer prometedor como base para una representación física de la realidad.

De hecho, esta idea resultaba atractiva por más de una razón. Irónicamente, uno de los motivos más concretos — relacionado con la física experimental de las interacciones entre hadrones— parece haber sido obviado por completo por los desarrollos más modernos en teoría de cuerdas, y no estoy seguro de que conserve una buena posición dentro de este campo, más allá de su valor histórico. Pero debo analizarlo igualmente (y lo haré, sobre todo, en §1.6), junto con algunos otros de los elementos de los antecedentes de la física fundamental de partículas sobre los que se erigen los principios básicos de la teoría de cuerdas.

Primero, permítanme explicar lo que es un hadrón. Recordemos que un átomo ordinario está formado por un núcleo con carga positiva, rodeado de electrones de carga negativa que orbitan a su alrededor. El núcleo está compuesto de protones y neutrones, denominados conjuntamente nucleones (N); cada protón tiene una carga eléctrica de una unidad positiva (la unidad de carga ha sido elegida de forma que el electrón tiene una unidad negativa) y cada neutrón, una carga eléctrica nula. La fuerza eléctrica atractiva entre cargas positivas y negativas es lo que mantiene a los electrones, de carga negativa, orbitando alrededor del núcleo, cargado positivamente. Pero, si las fuerzas eléctricas fuesen las únicas relevantes, el propio núcleo (aparte del de hidrógeno, que posee un solo protón) estallaría en sus varios componentes, porque los protones se repelerían entre sí, ya que todos ellos poseen carga positiva.

Por lo tanto, debe haber otra fuerza, más intensa, que mantiene el núcleo unido, que se conoce como fuerza (nuclear) fuerte. Además, existe algo llamado fuerza (nuclear) débil, de especial relevancia en relación con la desintegración nuclear, pero que no constituye el componente principal de las fuerzas entre nucleones. Más adelante diré algo sobre la fuerza débil.

No todas las partículas se ven afectadas de manera directa por la fuerza fuerte —por ejemplo, a los electrones no les influye—; las que sí la sienten son las partículas comparativamente masivas llamadas hadrones (del griego hadros, que significa «denso»). Así pues, los protones y los neutrones son ejemplos de hadrones, pero ahora sabemos que existen muchos otros tipos. Entre ellos están los primos de protones y neutrones llamados bariones (de barys, que significa «pesado»), que, además de los propios neutrones y protones, incluye los lambda ( $\Lambda$ ), sigma ( $\Sigma$ ), xi ( $\Xi$ ), delta ( $\Delta$ ) y omega ( $\Omega$ ), la mayoría de los cuales existen en distintas versiones con diferentes valores de la carga eléctrica, así como en una serie de versiones excitadas (que giran más rápido sobre sí mismos). Todas estas partículas son más masivas que el protón y el neutrón. La razón por la que no vemos que estas partículas más exóticas formen parte de los átomos ordinarios es que son muy inestables y se desintegran rápidamente, hasta acabar en forma de protones y neutrones, deshaciéndose de su exceso de masa en forma de energía (de acuerdo con la famosa  $E = mc^2$  de Einstein). El protón, por su parte, posee la masa de alrededor de 1.836 electrones y el neutrón, de unos 1.839 electrones. Entre los bariones y los electrones existe otra clase de hadrones, llamados mesones, cuyos integrantes más conocidos son el pión ( $\pi$ ) y el kaón ( $K$ ), de cada uno de los cuales existe una

versión con carga ( $\mu^+$  y  $\mu^-$ , cada uno de los cuales posee una masa de alrededor de 273 electrones;  $K^+$  y  $K^-$ , con una masa de unos 966 electrones cada uno) y otra sin ella ( $\mu^0$ , que posee una masa de alrededor de 264 electrones;  $K^0$  y  $\bar{K}^0$ , cada uno de los cuales tiene una masa de unos 974 electrones). Lo habitual aquí es usar una barra sobre el símbolo de la partícula para denotar la antipartícula; no obstante, tengamos en cuenta que los antipiones son también piones, mientras que un antikaón es distinto de un kaón. De nuevo, estas partículas tienen muchos primos y versiones excitadas (con mayor velocidad de giro).

Empezamos a ver que todo esto es muy complicado y dista mucho de cómo era en los vertiginosos días de principios del siglo XX, cuando parecía que el protón, el neutrón y el electrón (y dos o tres partículas sin masa, como el fotón, la partícula de la luz) eran, más o menos, todo lo que había. Con el paso de los años las cosas se fueron complicando, hasta que al final, entre aproximadamente 1970 y 1973, tomó forma una visión unificada de todo ello, conocida como modelo estándar de la física de partículas [Zee, 2010; Thomson, 2013]. Según este esquema, todos los hadrones están compuestos de quarks y/o sus antipartículas, los antiquarks. En la actualidad, se entiende que cada barión está compuesto de tres quarks y cada mesón (ordinario), de un quark y un antiquark. Existen seis sabores distintos de quark, conocidos (de una forma bastante extraña y poco imaginativa) como arriba, abajo, encanto, extraño, cima y fondo, que poseen, respectivamente, las cargas eléctricas  $\frac{2}{3}$ ,  $-\frac{1}{3}$ ,  $\frac{1}{3}$ ,  $-\frac{2}{3}$ ,  $\frac{1}{3}$  y  $-\frac{1}{3}$ . En un primer momento, los valores fraccionales de la carga resultan claramente extraños, pero para las partículas libres

observadas (como los bariones y los mesones) la carga eléctrica siempre posee un valor entero.

El modelo estándar no solo sistematiza la colección de las partículas fundamentales de la naturaleza, sino que también proporciona una buena descripción de las fuerzas que les afectan. Tanto la fuerza fuerte como la débil se describen en función de un elegante procedimiento matemático — conocido como teoría gauge (o de recalibración)— que hace un uso crucial de la idea de un fibrado, de la que se ofrece una breve descripción en §A.7, y a la que volveré, en particular en §1.8. El espacio base del fibrado (concepto que se explica en §A.7) es el espaciotiempo y, en el caso de la fuerza fuerte (que es el más transparente desde un punto de vista matemático), la fibra  $F$  se describe en función de un concepto conocido como color, que se asigna a cada quark (para cada quark hay disponibles tres colores alternativos). Por este motivo, la teoría física de la interacción fuerte se conoce como cromodinámica cuántica (QCD, por sus siglas en inglés). No tengo la intención de adentrarme aquí en un análisis en profundidad de la QCD, porque es difícil describirla adecuadamente sin utilizar más matemáticas que las que puedo incluir aquí [véanse Tsou y Chan, 1993; Zee, 2003]. Además, «no está de moda», en el sentido en que utilizo la expresión aquí, porque las ideas, aunque puedan parecer exóticas y extrañas, de hecho funcionan extraordinariamente bien, no solo porque constituyen un formalismo matemático coherente y muy cohesionado, sino porque obtienen una confirmación excelente en los resultados experimentales. El sistema de la QCD se estudia en cualquier departamento de investigación en el campo de la física dedicado en serio a la teoría de las interacciones fuertes, pero no está simplemente de moda, en el sentido

que esta expresión tiene aquí, porque se estudia tan a fondo por motivos científicos de peso.

No obstante, a pesar de todas sus virtudes, hay también poderosas razones científicas para tratar de ir más allá del modelo estándar. Una de ellas es que en él existen alrededor de treinta números para los que la teoría no ofrece explicación alguna. Entre ellos están cosas como las masas de los quarks y los leptones, magnitudes conocidas como parámetros de acoplamiento entre fermiones (como el ángulo de Cabibbo), el ángulo de Weinberg, el ángulo zeta, los acoplamientos gauge y parámetros relacionados con el mecanismo de Higgs. Hay otro importante inconveniente relacionado con este asunto, ya presente en los sistemas que existían con anterioridad a la aparición del modelo estándar, y que este resuelve solo en parte. Se trata del inquietante problema de los valores infinitos (resultados absurdos que se obtienen a partir de expresiones divergentes, como las que se muestran en §A.10) que surgen en la teoría cuántica de campos (QFT, por sus siglas en inglés), la forma de la mecánica cuántica esencial no solo para la QCD y otros aspectos del modelo estándar, sino para todas las aproximaciones modernas a la física de partículas, así como para muchos aspectos de la física básica.

Tendré mucho más que decir sobre la mecánica cuántica en general en el capítulo 2. De momento, limitemos nuestra atención a una característica muy específica pero fundamental de la mecánica cuántica, que puede considerarse una raíz del problema de los valores infinitos en la QFT, y veremos también que el método convencional de tratar con estos infinitos impide dar una respuesta completa a la cuestión de derivar los treinta y tantos números que carecen de explicación en el modelo estándar. La teoría de

cuerdas viene motivada en gran medida por una ingeniosa propuesta para esquivar los valores infinitos de la QFT, como veremos en §1.6, y parece ofrecer así alguna esperanza de encontrar la vía para resolver el misterio de los números inexplicados.

#### **1.4.EL PRINCIPIO DE SUPERPOSICIÓN EN LA QFT**

Una piedra angular de la mecánica cuántica es el principio de superposición, una característica común a toda la teoría cuántica, no solo la QFT. En particular, resultará esencial para las discusiones críticas del capítulo 2. En este, para arrojar algo de luz sobre la fuente del problema de los infinitos en la QFT, tendré que introducir brevemente este principio aquí, aunque el grueso de la discusión acerca de la mecánica cuántica tendrá lugar en el siguiente capítulo (véanse, en particular, §§2.5 y 2.7).

Para poner de manifiesto el papel que el principio de superposición desempeña en la QFT, consideremos situaciones del siguiente tipo. Supongamos que tenemos un proceso físico que lleva a un determinado resultado observado. Supondremos que a este resultado se podría haber llegado a través de una acción intermedia  $\Psi$ , pero que existe también otra posible acción intermedia  $\Phi$  que podría dar lugar básicamente al mismo resultado observado. Entonces, según el principio de superposición, debemos considerar que, en un sentido apropiado, tanto  $\Psi$  como  $\Phi$  podrían haber tenido lugar concurrentemente como acción intermedia. Esto, a todas luces, choca de manera frontal con nuestra intuición, puesto que a escala macroscópica ordinaria no vemos que posibilidades alternativas distintas tengan lugar al mismo tiempo. Pero en los eventos microscópicos, donde no tenemos la posibilidad de observar

directamente si ha ocurrido una actividad intermedia u otra, debemos suponer que ambas podrían haberse producido juntas, en lo que se conoce como una superposición cuántica.

El ejemplo paradigmático de este tipo de fenómeno ocurre en el famoso experimento de la doble rendija, que suele usarse en las introducciones a la mecánica cuántica. Aquí consideramos una situación en la que un haz de partículas (por ejemplo, electrones o fotones) se dirige contra una pantalla, de tal forma que, en su recorrido desde la fuente hasta la pantalla, el haz debe atravesar un par de rendijas paralelas y próximas entre sí (véase la figura 1-2(a)). En la situación que estamos considerando, al llegar a la pantalla cada partícula deja una clara marca oscura en una posición concreta sobre la misma, indicativa de la propia naturaleza corpuscular de la partícula. Pero, a medida que van llegando muchas de estas partículas, se va creando un patrón de interferencia con franjas claras y oscuras; estas últimas se producen allí donde muchas partículas chocan con la pantalla, mientras que las primeras ocurren donde la alcanzan relativamente pocas partículas (figura 1-2(d)). Un minucioso análisis estándar de la situación[2] permite llegar a la conclusión de que cada partícula cuántica debe, en algún sentido, atravesar ambas rendijas al mismo tiempo, a la manera de una extraña clase de superposición de los dos caminos alternativos que podría seguir.

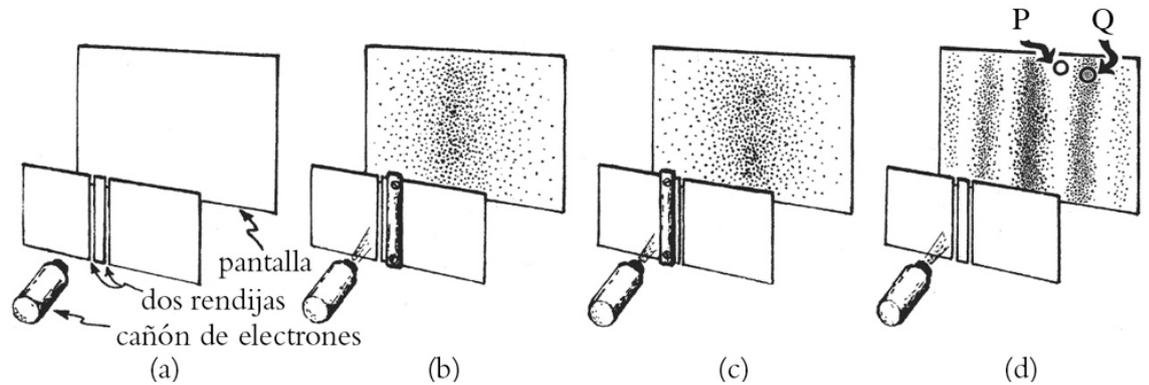


FIGURA 1-2. El experimento de la doble rendija. Se lanzan los electrones hacia una pantalla a través de un par de rendijas muy próximas entre sí (a). Si solo una de ellas está abierta (b), (c), en la pantalla se registra un patrón de impactos de aspecto aleatorio, dispersos alrededor del camino directo a través de la rendija. Sin embargo, si ambas rendijas están abiertas (d), el patrón adquiere una apariencia de franjas, y ahora hay algunos lugares (por ejemplo, P) a los que no puede llegar la partícula, aunque sí podría hacerlo si solo estuviese abierta una de las rendijas; además, en otros sitios (por ejemplo, Q) la intensidad de recepción es cuatro veces superior a la correspondiente a una sola rendija.

El razonamiento en que se basa tan extraña conclusión deriva del hecho de que, si se cubre cualquiera de las rendijas mientras se deja la otra abierta (figura 1-2(b) y (c)), no aparecen las franjas sino solo una iluminación bastante uniforme que es más oscura en el centro. Sin embargo, cuando ambas rendijas están abiertas, entre las franjas más oscuras aparecen regiones más claras en la pantalla, en lugares que son perfectamente oscuros cuando solo está abierta una de las rendijas. De alguna manera, cuando la partícula tiene a su alcance ambos caminos, esos lugares más claros se atenúan, mientras que los lugares oscuros se

acentúan. Si cada partícula hiciese solo lo que podría hacer cuando solo una de las rendijas está abierta, o lo que podría hacer cuando está abierta la otra, los efectos de los caminos simplemente se sumarían, y no obtendríamos estas extrañas franjas de interferencia. Esto sucede únicamente porque los dos caminos posibles están disponibles para la partícula, que explora ambas alternativas para dar lugar al efecto en cuestión. En cierto sentido, estos recorridos coexisten para la partícula cuando se encuentra entre la fuente y la pantalla.

Evidentemente, esto no se corresponde en absoluto con nuestra experiencia del comportamiento de los objetos macroscópicos. Por ejemplo, si dos habitaciones están conectadas entre sí por dos puertas diferentes y se observa que un gato estaba primero en una habitación y después en la otra, normalmente inferiríamos que ha pasado por una u otra de las puertas, no que podría, de alguna manera extraña, haber atravesado ambas puertas a la vez. Pero, con un objeto del tamaño de un gato, sería posible, sin perturbar significativamente sus acciones, tomar mediciones continuas de su posición, y de esta manera determinar por cuál de las dos puertas pasó de hecho. Si lográsemos hacer algo similar para una única partícula cuántica en el experimento de las dos rendijas descrito antes, tendríamos que perturbar su comportamiento hasta tal punto que daría lugar a la destrucción del patrón de interferencia en la pantalla. El comportamiento ondulatorio de una partícula cuántica concreta que da lugar a las franjas de interferencia claras y oscuras en la pantalla depende del hecho de que no somos capaces de determinar cuál de las dos rendijas atravesó, lo que abre la posibilidad de este desconcertante estado intermedio superpuesto de la partícula.

En este experimento de las dos rendijas, podemos ver la absoluta rareza del comportamiento de las diferentes partículas cuánticas fijándonos muy especialmente en un punto P de la pantalla situado en mitad del hueco entre las franjas oscuras, pues observamos que la partícula simplemente es incapaz de llegar a P cuando ambas rendijas están abiertas, mientras que, cuando solo una lo está, la partícula llega sin problemas a P a través de ella. Cuando ambas rendijas están abiertas, las dos posibilidades de que dispone la partícula para alcanzar P de alguna forma se cancelan mutuamente. Sin embargo, en otro lugar de la pantalla —por ejemplo, Q—, donde el patrón de interferencia es lo más oscuro posible, vemos que, en lugar de cancelarse, los dos caminos posibles parecen reforzarse entre sí, de manera que, cuando las dos rendijas están abiertas, la probabilidad de que la partícula llegue a P es cuatro veces mayor de lo que lo sería si solo estuviese abierta una de ellas, no solo el doble, como sucedería si tuviésemos un objeto clásico ordinario en lugar de una partícula cuántica; véase la figura 1-2(d). Estas extrañas características son consecuencia de la conocida como regla de Born, que relaciona las intensidades en las superposiciones con las probabilidades reales de ocurrencia, como veremos en breve.

La palabra «clásico», dicho sea de paso, cuando se emplea en el contexto de las teorías, modelos o situaciones, significa simplemente «no cuántico». En particular, la teoría general de la relatividad de Einstein es una teoría clásica, a pesar de haber sido propuesta después de que se hubiesen desarrollado muchas de las ideas fundamentales de la teoría cuántica (como la del átomo de Bohr). De manera muy particular, los sistemas clásicos no están sujetos a las curiosas superposiciones de posibilidades alternativas que

acabamos de ver y que caracterizan al comportamiento cuántico, como explicaré brevemente a continuación.

Dejaré el análisis completo de los fundamentos de nuestra comprensión actual de la física cuántica para el capítulo 2 (véase, en particular, de §2.3 en adelante). De momento, recomiendo que nos limitemos a aceptar la extraña regla matemática según la cual la mecánica cuántica moderna describe tales estados intermedios, que resulta ser extraordinariamente precisa. Pero ¿cuál es esta extraña regla? El formalismo cuántico afirma que semejante estado intermedio superpuesto, cuando solo existen dos posibilidades intermedias alternativas  $\Psi$  y  $\Phi$ , se expresa matemáticamente como una especie de suma  $\Psi + \Phi$  de las posibilidades, o, de forma más general, como una combinación lineal (véanse §§A.4 y A.5),

$$w\Psi + z\Phi,$$

donde  $w$  y  $z$  son números complejos (en los que aparece  $i = \sqrt{-1}$ , como se describe en §A.9) que no se anulan simultáneamente. Además, nos veremos obligados a considerar que estas superposiciones complejas de estados deben poder persistir en un sistema cuántico hasta el instante mismo en que el sistema sea observado, momento en que la superposición de alternativas debe sustituirse por una combinación probabilística de dichas alternativas. Esto es sin duda extraño, pero en §§2.5 a 2.7 y 2.9 veremos cómo usar estos números complejos —que suelen denominarse amplitudes— y cómo se relacionan, de maneras a veces sorprendentes, con las probabilidades, así como con la evolución temporal de los sistemas físicos a escala cuántica (ecuación de Schrödinger); también tienen relación, en un sentido fundamental, con el sutil comportamiento del espín

de una partícula cuántica, e incluso con la 3-dimensionalidad del espacio físico ordinario. Aunque las conexiones precisas entre estas amplitudes y las probabilidades (la regla de Born) no se abordarán plenamente en este capítulo (ya que para ello necesitamos los conceptos de ortogonalidad y normalización para los  $\Psi$  y  $\Phi$ , que es preferible dejar para §2.8), lo esencial de la regla de Born se expone a continuación.

Una medición, dirigida a determinar si un sistema se encuentra en un estado  $\Psi$  o en un estado  $\Phi$ , cuando se encuentra con el estado superpuesto  $w\Psi + z\Phi$ , obtiene:

cociente entre la probabilidad de  $\Psi$  y la probabilidad de  $\Phi$  =  
cociente entre  $|w|^2$  y  $|z|^2$ .

Tengamos en cuenta (véanse §§A.9 y A.10) que el cuadrado del módulo  $|z|^2$  de un número complejo  $z$  es la suma de los cuadrados de sus partes real e imaginaria, igual también al cuadrado de la distancia desde el origen del plano de Wessel hasta  $z$  (figura A-42 en §A.10). Cabe señalar asimismo que el hecho de que las probabilidades se obtengan a partir de los cuadrados de los módulos de las amplitudes explica la cuadruplicación de la intensidad que hemos visto antes, cuando las contribuciones se refuerzan mutuamente en el experimento de la doble rendija (véase también la parte final de §2.6).

Debemos tener muy claro que la noción de más en estas superposiciones es muy distinta del concepto ordinario de y (a pesar de que, en el habla corriente, hoy en día es habitual usar más para referirse a y), o incluso de o. Lo que esto quiere decir es que, en cierto sentido, debemos entender que las dos posibilidades se suman de alguna manera abstracta. Así, en el caso del experimento de la doble rendija, donde  $\Psi$  y  $\Phi$

representan dos posiciones transitorias diferentes de una sola partícula,  $\Psi + \Phi$  no representa dos partículas, una en cada posición («una partícula en la posición  $\Psi$  y una partícula en la posición  $\Phi$ », lo que implicaría dos partículas en total), ni tampoco debemos entender que ambas son solo alternativas ordinarias, una u otra de las cuales sucedió efectivamente, aunque no sabemos cuál. Debemos pensar en una única partícula que de algún modo ocupa ambas posiciones al mismo tiempo, superpuestas de acuerdo con esta extraña operación «más» mecanocuántica. Desde luego, esto parece sumamente raro, y los físicos de principios del siglo XX no se habrían planteado considerar algo así sin tener muy buenos motivos para ello. En el capítulo 2 analizaré algunos de estos motivos; de momento solo le pido al lector que acepte que este formalismo en efecto funciona.

Es importante advertir que, según la mecánica cuántica estándar, a este procedimiento de superposición se lo considera universal y, en consecuencia, válido también si existen más de solo dos alternativas para el estado intermedio. Por ejemplo, si hay tres posibilidades alternativas,  $\Psi$ ,  $\Phi$  y  $\Gamma$ , tendremos que considerar superposiciones triples como  $w\Psi + z\Phi + u\Gamma$  (donde  $w$ ,  $z$  y  $u$  son números complejos, no todos ellos nulos). Por consiguiente, si hubiese cuatro estados intermedios alternativos, deberíamos considerar superposiciones cuádruples, y así sucesivamente. La mecánica cuántica lo exige, y existe un excelente soporte experimental para este comportamiento a la escala submicroscópica de la actividad cuántica. Es extraño, qué duda cabe, pero las matemáticas son sólidas. Estas matemáticas son, hasta ahora, tan solo las de un espacio vectorial, con números escalares complejos, como las que se exponen en §§A.3, A.4, A.9 y A.10; a partir de §2.3 veremos

más sobre el papel ubicuo de las superposiciones cuánticas. Sin embargo, las cosas son considerablemente más complicadas en la QFT, porque con frecuencia debemos considerar situaciones en las que las posibilidades intermedias son infinitas. En consecuencia, nos vemos abocados a considerar sumas infinitas de alternativas, y se cierne sobre nosotros la posibilidad de que una de estas sumas se traduzca en una serie cuya suma diverja al infinito (de manera similar a como se muestra en §§A.10 y A.11).

### **1.5. EL PODER DE LOS DIAGRAMAS DE FEYNMAN**

Tratemos de entender un poco más en detalle cómo se producen realmente esas divergencias. En física de partículas, lo que tenemos que considerar son situaciones en las que varias partículas se unen para crear otras partículas, otras situaciones en las que algunas de las partículas podrían dividirse para dar lugar a más partículas, y otras aún en las que pares de estas partículas podrían volver a unirse, etc., etc., lo que puede acabar dando lugar a procesos muy complicados. Los tipos de situación que suelen interesar a los físicos de partículas involucran a un determinado conjunto de partículas que se juntan —a menudo a velocidades relativas cercanas a la de la luz—, y esta combinación de colisiones y separaciones resulta en otro conjunto de partículas que surge como resultado de todo ello. El proceso total implicaría una inmensa superposición cuántica de todos los diferentes tipos posibles de procesos intermedios que podrían tener lugar y son compatibles con las situaciones inicial y final. Un ejemplo de un proceso complicado de este estilo se ilustra en el diagrama de Feynman de la figura 1-3.

No nos equivocamos mucho si vemos un diagrama de Feynman como un diagrama espaciotemporal de un conjunto particular de procesos de partículas. Como me dedico a la teoría de la relatividad, en lugar de ser un físico profesional de partículas o un experto en la QFT, me gusta representar el tiempo de forma que aumente hacia la parte superior de la página; normalmente, esos profesionales prefieren que el tiempo avance de izquierda a derecha. Los diagramas de Feynman (o gráficas de Feynman) llevan el nombre del destacado físico estadounidense Richard Phillips Feynman. En la figura 1-4 se pueden ver varios diagramas básicos de este tipo. La figura 1-4(a) muestra la división de una partícula en dos, mientras que la figura 1-4(b) representa la combinación de dos partículas para crear una tercera.

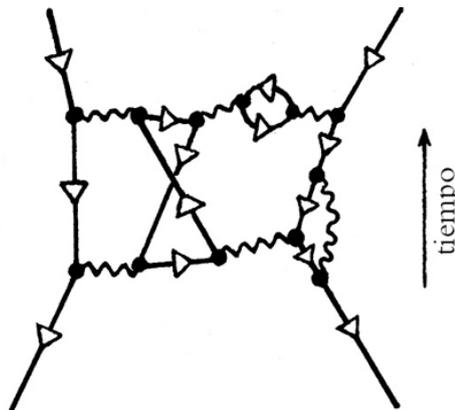


FIGURA 1-3. Un diagrama de Feynman (dibujado de manera que el tiempo avanza hacia arriba) es una representación espaciotemporal esquemática (con una interpretación matemática bien definida) de un proceso de partículas que suele implicar la creación, destrucción e intercambios de partículas intermedias. Las líneas onduladas representan fotones. Las flechas triangulares denotan carga eléctrica

(positiva, si la flecha apunta hacia arriba; negativa, si lo hace hacia abajo).

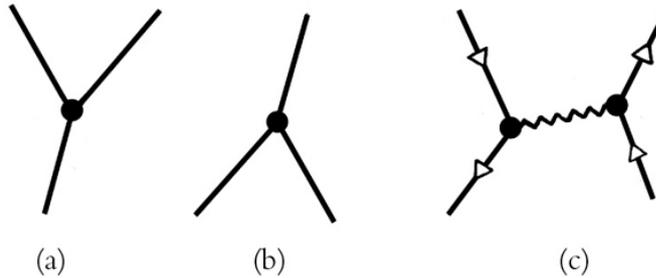


FIGURA 1-4. Diagramas de Feynman elementales: (a) la partícula se divide en dos; (b) dos partículas se combinan para crear otra partícula; (c) dos partículas con cargas opuestas (por ejemplo, un electrón y un positrón) «intercambian» un fotón.

En la figura 1-4(c), vemos el intercambio de una partícula (por ejemplo, un fotón, el cuanto de campo electromagnético o luz, que se representa mediante la línea ondulada) entre dos partículas. El uso del término «intercambio» para este proceso, aunque habitual entre físicos de partículas, quizá aquí resulte un poco raro, ya que un único fotón solo pasa de una partícula externa a otra, aunque de una manera que (deliberadamente) no deja claro qué partícula es la emisora y cuál la receptora. El fotón que interviene en este intercambio suele calificarse de virtual y su velocidad no está constreñida a cumplir con los requisitos de la relatividad. Puede que el uso coloquial habitual del término «intercambio»