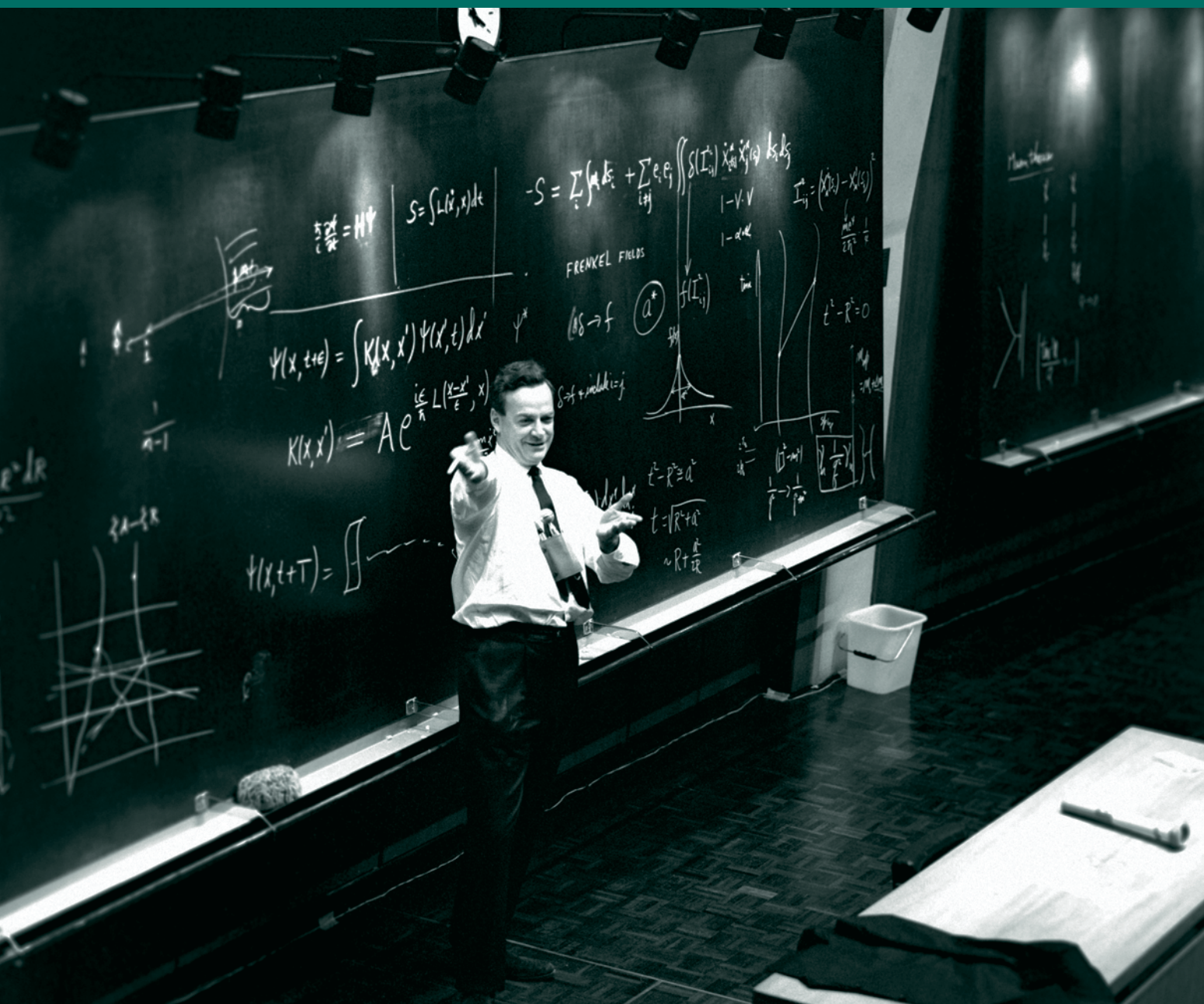


Lecciones de física de Feynman

III. Mecánica cuántica

Richard P. Feynman, Robert B. Leighton
y Matthew Sands



EDICIONES
CIENTÍFICAS
UNIVERSITARIAS

TEXTO CIENTÍFICO
UNIVERSITARIO

Richard P. Feynman (Nueva York, 1918-Los Ángeles, 1988), reconocido físico estadounidense, estudió ciencias en el Instituto Tecnológico de Massachusetts (MIT) y el doctorado en la Universidad de Princeton. Enseñó en Cornell y en el Instituto Tecnológico de California (Caltech). En 1965 recibió el Premio Nobel de Física, junto con Sin-Itiro Tomonaga y Julian Schwinger por su trabajo fundamental en electrodinámica cuántica, con profundas consecuencias para la física de las partículas elementales.

Robert B. Leighton (Detroit, 1919-Pasadena, 1997) realizó un trabajo pionero en física de estado sólido, física de rayos cósmicos, los comienzos de la física moderna de partículas, física solar, fotografía planetaria, astronomía infrarroja y astronomía de ondas milimétricas y submilimétricas a lo largo de su vida. Fue ampliamente conocido por su diseño innovador de instrumentos científicos, y muy admirado como profesor.

Matthew Sands (Oxford, Massachusetts 1919-Santa Cruz, California, 2014) fue un reconocido físico y profesor, doctor por el MIT. Después de la guerra, Sands ayudó a fundar la Federación de Científicos Atómicos de Los Álamos, que presionó contra el uso de armas nucleares. Participó en la construcción del sincrotrón de Caltech y fue director del Acelerador Lineal de Stanford, entre otros cargos científicos y educativos.

EDICIONES CIENTÍFICAS UNIVERSITARIAS

SERIE TEXTO CIENTÍFICO UNIVERSITARIO

LECCIONES DE FÍSICA DE FEYNMAN, III

RICHARD P. FEYNMAN
ROBERT B. LEIGHTON • MATTHEW SANDS

Lecciones de física de Feynman

III. Mecánica cuántica



FONDO DE CULTURA ECONÓMICA

Primera edición en inglés, 1965
Primera edición en español, 1971
Primera edición, FCE, 2022

[Primera edición en libro electrónico, 2023]

Feynman, Richard P., Robert B. Leighton y Matthew Sands

Lecciones de física de Feynman III. Mecánica cuántica / Richard P. Feynman, Robert B. Leighton, Matthew Sands ; trad. de Carlos Alberto Heras, Juan Martín y Marfil ; colab. de Ricardo Gómez. — México : FCE, 2022

[416] p. ; ilus. ; 28 × 21 cm — (Colec. Ediciones Científicas Universitarias)

Título original: *The Feynman Lectures on Physics, III. Quantum mechanics*

ISBN 978-607-16-5972-9 (Obra completa)

ISBN 978-607-16-7574-3 (Vol. 3)

1. Mecánica cuántica 2. Física I. Leighton, Robert B., coaut. II. Sands, Matthew, coaut. III. Heras, Carlos Alberto, tr. IV. Martín y Marfil, Juan, tr. V. Gómez, Ricardo, colab. VI. Ser. VII. t.

LC QC23.F47

Dewey 530 F574l Vol. 3

Distribución mundial

La presente edición se basa en la *New Millennium Edition* publicada por Basic Books e incluye la corrección de erratas hasta su 7ª reimpresión. Ésta es una versión corregida y revisada de la traducción de Carlos Alberto Heras y Juan Martín y Marfil, con la colaboración de Ricardo Gómez, que fue publicada originalmente bajo el sello del Fondo Educativo Interamericano, S. A.

© 1965, 2006, 2010, por el California Institute of Technology,

Michael A. Gottlieb y Rudolf Pfeiffer

Título original: *The Feynman Lectures on Physics,*

III. Quantum mechanics

Proyecto apoyado por el Consejo Nacional de Ciencia y Tecnología



D. R. © 2022, Fondo de Cultura Económica

Carretera Picacho-Ajusco, 227; 14110 Ciudad de México

www.fondodeculturaeconomica.com

Comentarios: editorial@fondodeculturaeconomica.com

Tel.: 55-5227-4672

Diseño de portada: Laura Esponda Aguilar

Fotografía de portada: Conferencia post-Nobel de Richard Feynman en el CERN, 1965.

Crédito: CERN/Science Photo Library

Se prohíbe la reproducción total o parcial de esta obra, sea cual fuere el medio, sin la anuencia por escrito del titular de los derechos.

ISBN 978-607-16-5972-9 (rústico, Obra completa)

ISBN 978-607-16-7574-3 (rústico, Vol. 3)

ISBN 978-607-16-6120-3 (electrónico-pdf, Obra completa)

ISBN 978-607-16-7792-1 (electrónico-pdf, Vol. 3)

Impreso en México • *Printed in Mexico*

Sobre Richard Feynman

Nacido en 1918 en la ciudad de Nueva York, Richard P. Feynman recibió su doctorado de la Universidad de Princeton en 1942. A pesar de que era muy joven, tuvo un papel importante en el Proyecto Manhattan en Los Álamos durante la segunda Guerra Mundial. Más tarde, dio clases en la Universidad Cornell y en el Instituto de Tecnología de California. En 1965 recibió el premio Nobel de física, junto con Sin-Itiro Tomonaga y Julian Schwinger, por su trabajo en la electrodinámica cuántica.

El doctor Feynman ganó el premio Nobel por haber resuelto con éxito algunos problemas de la teoría de la electrodinámica cuántica. También creó una teoría matemática que explica el fenómeno de la superfluidez en el helio líquido. Después, con Murray Gell-Mann, realizó un trabajo fundamental en el área de las interacciones débiles como la desintegración beta. En años posteriores, Feynman desempeñó un papel clave en el desarrollo de la teoría de los quarks al presentar su modelo de partones de procesos de colisión de protones de alta energía.

Más allá de estos logros, el doctor Feynman introdujo nuevas técnicas básicas de computación y notación en la física: se destacan los omnipresentes diagramas de Feynman que, quizá más que ningún otro formalismo en la historia científica cercana, han cambiado la manera en la que los procesos físicos básicos son conceptualizados y calculados.

Feynman fue un maestro sobresaliente en su labor. De todos sus numerosos galardones, estaba orgulloso en particular de la Medalla Oersted para la enseñanza, a la cual se hizo acreedor en 1972. Un reseñista de *Scientific American* describió las *Lecciones de física de Feynman*, publicadas originalmente en 1963, como “arduas, pero sustanciosas y llenas de sabor. Después de 25 años son la guía para maestros y para los mejores alumnos que comienzan”. Con el fin de mejorar la comprensión de la física entre los no especialistas, el doctor Feynman escribió: *El carácter de la ley física y Electrodinámica cuántica. La extraña teoría de la luz y la materia*. También fue autor de varias publicaciones avanzadas que se han vuelto referencias y manuales clásicos para investigadores y estudiantes.

Richard Feynman fue un hombre público productivo. Su trabajo en la comisión Challenger es perfectamente conocido, sobre todo su famosa demostración de la susceptibilidad de las juntas tóricas al frío, un experimento elegante que requería de nada más que un vaso de agua helada y una abrazadera. Fueron menos célebres los esfuerzos del doctor Feynman en los años sesenta en el comité curricular de la Universidad Estatal de California, donde protestó por la mediocridad de los libros de texto.

Una recitación de los incontables logros científicos y educativos de Richard Feynman no puede capturar adecuadamente la esencia de este hombre. Como lo sabe cualquier lector de sus publicaciones, incluso las más técnicas, la personalidad dinámica y multifacética de Feynman resplandece a lo largo de toda su obra. Además de físico, en muchas ocasiones fue reparador de radios, forzador de cerraduras, artista, bailarín, intérprete de bongó y descifrador de jeroglíficos mayas. Con una curiosidad perpetua por su mundo, fue un empirista ejemplar.

Richard Feynman murió el 15 de febrero de 1988 en Los Ángeles.

Prefacio a la Edición del nuevo milenio

Han pasado casi 50 años desde que Richard Feynman impartió el curso introductorio de física en el Instituto de Tecnología de California (Caltech) que dio origen a estos tres volúmenes, las *Lecciones de física de Feynman*. En esos 50 años, nuestra comprensión del mundo físico ha cambiado en gran medida, pero esta obra ha perdurado. Las lecciones de Feynman son tan poderosas hoy en día como lo eran recién publicadas, gracias a lo singular de sus observaciones sobre la física y su método de enseñanza. Físicos principiantes y experimentados las han estudiado por igual en todo el mundo; han sido traducidas al menos a una docena de lenguas, con más de 1.5 millones de ejemplares impresos sólo en inglés. Quizá ninguna otra serie de libros de física haya tenido un impacto tan amplio y duradero.

Esta *Edición del nuevo milenio* abre paso a una nueva era para las *Lecciones de física de Feynman*: la era de la edición digital en el siglo XXI. Las *Lecciones* se han vuelto *Lecciones-e* cuyo texto y ecuaciones se expresan en el lenguaje de composición tipográfica L^AT_EX y cuyas figuras fueron hechas de nuevo en su totalidad con software de dibujo moderno.

Las consecuencias para la edición *impresa* de esta edición *no* sorprenden; se ve casi igual a los libros rojos originales que los estudiantes de física han conocido y disfrutado por décadas. Las diferencias principales están en un índice alfabético ampliado y mejorado, en la corrección de 885 erratas que los lectores encontraron a lo largo de cinco años desde la primera impresión de la edición anterior, y en la facilidad de corregir erratas que los futuros lectores lleguen a encontrar. Regresaré a esto líneas adelante.

La *eBook Version* de esta edición y la *Enhanced Electronic Version* son innovaciones electrónicas. En contraste con la mayoría de las versiones electrónicas de libros técnicos del siglo XX, cuyas ecuaciones, figuras y en ocasiones incluso texto pierden calidad en cuanto uno trata de ampliarlos, el manuscrito en L^AT_EX de la *Edición del nuevo milenio* hace posible la creación de libros electrónicos de la más alta calidad, en los que todas las características de la página (excepto las fotografías) pueden ampliarse sin límites y conservar sus formas y definición precisas. Y la *Enhanced Electronic Version*, que cuenta con audio y fotografías de pizarrones de las lecciones originales de Feynman, además de hipervínculos a otros recursos, es una innovación que le habría dado mucha satisfacción a Feynman.

Recuerdos de las lecciones de Feynman

Estos tres volúmenes son un tratado pedagógico autosuficiente. También son un registro histórico de las lecciones universitarias de física de Feynman entre 1961 y 1964, una materia obligatoria para todos los estudiantes de primer y segundo años en Caltech, sin importar la especialidad a la que aspirarán.

Los lectores se preguntarán, como lo hice yo, cuál fue el impacto que las lecciones de Feynman tuvieron en los estudiantes que las presenciaron. En su Prefacio a estos volúmenes, Feynman ofreció un punto de vista algo negativo al respecto: “No creo que me haya ido muy bien con los estudiantes”, escribió. Matthew Sands, en su breve texto autobiográfico en *Feynman’s Tips on Physics*, expresó un punto de vista mucho más optimista. Por pura curiosidad, en la primavera de 2005 envié correos electrónicos o hablé con una muestra *cuasi* aleatoria de 17 alumnos (de unos 150) de la clase de Feynman entre 1961 y

1963 —entre ellos algunos a los que se les dificultó mucho la clase y otros que no tuvieron ningún problema para dominarla, biólogos, químicos, ingenieros, geólogos, matemáticos y astrónomos, al igual que físicos—.

Puede que los años intermedios hayan esmaltado sus recuerdos de un matiz eufórico, pero cerca de 80 % de ellos recuerdan las lecciones de Feynman como uno de los puntos más memorables de sus años de universidad. “Era como ir a la iglesia.” Las lecciones eran “una experiencia transformativa”, “una experiencia de vida, probablemente lo más importante que obtuve de Caltech”. “Yo estudiaba biología, pero las lecciones de Feynman se destacan como uno de los puntos más altos en mi experiencia universitaria... aunque debo admitir que en esa época no podía hacer la tarea y casi no entregué ninguna.” “Yo me encontraba entre los estudiantes menos prometedores de su clase, y nunca me perdí una sola sesión... recuerdo y todavía puedo sentir el entusiasmo de Feynman por el descubrimiento... Sus lecciones tenían un... impacto emocional que quizá se perdió en las *Lecciones* impresas.”

En contraste, varios de los estudiantes tienen recuerdos negativos debido en gran medida a dos cuestiones: (i) “No podías aprender a trabajar con los problemas de tarea al asistir a las lecciones. Feynman era demasiado ingenioso: conocía trucos y sabía qué aproximaciones podían hacerse, tenía una intuición basada en la experiencia y en una genialidad que un estudiante que apenas comienza no posee”. Feynman y sus colegas, conscientes de esta falla en el curso, la resarcieron en parte con materiales que se han incorporado a *Feynman's Tips on Physics*: tres lecciones hechas por el propio Feynman sobre resolución de problemas, y un conjunto de ejercicios y respuestas preparado por Robert B. Leighton y Rochus Vogt. (ii) “La inseguridad de no saber lo que podría discutirse en la siguiente lección, la falta de un libro de texto o referencia que tuviera cualquier relación con el material de lectura y la consecuente inhabilidad que teníamos para adelantar lecturas eran muy frustrantes... Las lecciones me parecían emocionantes y comprensibles en el aula, pero afuera eran sánscrito (cuando trataba de reconstruir los detalles).” Desde luego, este problema se resolvió en estos tres volúmenes, la versión impresa de las *Lecciones de física de Feynman*. Éstas se volvieron el libro de texto que los estudiantes de Caltech han usado para estudiar durante muchos años después, y sobreviven hoy en día como uno de los mayores legados de Feynman.

Una historia de erratas

Feynman y sus coautores, Robert B. Leighton y Matthew Sands, produjeron las *Lecciones de física de Feynman* con gran rapidez, trabajando a partir de grabaciones y fotografías de pizarrones de las lecciones de Feynman[†] (ambas se incorporaron en la *Enhanced Electronic Version* de esta *Edición del nuevo milenio*). Debido al ritmo tan acelerado al que trabajaron Feynman, Leighton y Sands, era inevitable que muchos errores se colaran en la primera edición. Feynman acumuló largas listas de erratas alegadas en los años subsecuentes, erratas que hallaron los estudiantes y profesores de Caltech así como lectores alrededor del mundo. En los años sesenta e inicios de los setenta, Feynman hizo un espacio en su ajetreada vida para verificar la mayor parte de las erratas, aunque no todas, de los volúmenes I y II e insertar las correcciones en las reimpressiones siguientes. Pero el sentido del deber de Feynman nunca logró sobreponerse a su entusiasmo por nuevos descubrimientos para lograr que atendiera las erratas en el volumen III.[‡] Después de su muerte prematura en 1988, las listas de erratas de los tres volúmenes se guardaron en el Archivo de Caltech y quedaron ahí olvidadas.

[†] Para descripciones de la génesis de las lecciones de Feynman y de estos volúmenes, véanse el prefacio y las introducciones a cada uno de los tres volúmenes, e incluso véase el texto autobiográfico de Matt Sands en *Feynman's Tips on Physics*, y el prefacio especial a la *Edición conmemorativa* de las *Lecciones de Feynman*, escrito en 1989 por David Goodstein y Gerry Neugebauer, que también aparece en la *Edición definitiva* de 2005.

[‡] En 1975 comenzó a buscar las erratas del volumen III, pero se distrajo con otras cosas y nunca terminó la tarea, de modo que no se hicieron correcciones.

En 2002, Ralph Leighton (hijo del fallecido Robert Leighton y compañero de Feynman) me informó de las viejas erratas y de una nueva y larga lista que había compilado su amigo Michael Gottlieb. Leighton propuso que Caltech realizara una nueva edición de las *Lecciones de Feynman* con todas las erratas corregidas, y que la publicara junto con un nuevo volumen de material complementario, *Feynman's Tips on Physics*, que él y Gottlieb estaban preparando.

Feynman era mi héroe y fue mi amigo cercano. Cuando vi las listas de erratas y el contenido del nuevo volumen propuesto, accedí pronto a supervisar el proyecto en representación de Caltech (que durante mucho tiempo fue el hogar académico de Feynman, al cual él, Leighton y Sands habían encomendado todos los derechos y responsabilidades de las *Lecciones de Feynman*). Después de un año y medio de trabajo meticuloso por parte de Gottlieb, y del atento examen del doctor Michael Hartl (un investigador posdoctoral notable de Caltech que evaluó todas las erratas además del nuevo volumen), la *Definitive Edition of the Feynman Lectures on Physics* vio la luz en 2005; tenía alrededor de 200 erratas corregidas y estaba acompañada por *Feynman's Tips on Physics*, de Feynman, Gottlieb y Leighton.

Yo *pensé* que esa edición iba a ser la “definitiva”. Lo que no anticipé fue la respuesta entusiasta de los lectores alrededor del mundo ante el llamado de Gottlieb para identificar más erratas y enviarlas por medio del sitio web que él creó y aún mantiene, *The Feynman Lectures Website*, www.feynmanlectures.info. En los cinco años a partir de entonces, se han enviado 965 erratas nuevas que sobrevivieron el meticuloso escrutinio de Gottlieb, Hartl y Nate Bode (un estudiante de posgrado notable de Caltech, quien sucedió a Hartl como el examinador de erratas por parte de esta institución). De estas 965 erratas evaluadas, 80 se corrigieron en la cuarta reimpression de la *Definitive Edition* (agosto de 2006) y las otras 885 se corrigieron en esta primera impresión de la *Edición del nuevo milenio* (332 en el volumen I, 263 en el volumen II y 200 en el volumen III). Para los detalles de las erratas, véase www.feynmanlectures.info.

Está claro que liberar de errores las *Lecciones de física de Feynman* se ha vuelto una empresa comunitaria a nivel mundial. En nombre de Caltech agradezco a los 50 lectores que han contribuido desde 2005 y a los muchos más que lleguen a contribuir en los años por venir. Los nombres de todos los que contribuyeron se encuentran en www.feynmanlectures.info/flp_errata.html.

Casi todas las erratas han sido de tres tipos: (i) errores tipográficos en prosa; (ii) errores tipográficos y matemáticos en ecuaciones, tablas y figuras —errores en signos, números incorrectos (por ejemplo, un 5 que debería ser un 4), así como subíndices, signos de suma, paréntesis y términos faltantes en ecuaciones—; (iii) referencias cruzadas incorrectas a otros capítulos, tablas y figuras. Esta clase de errores, a pesar de que no son gravísimos para un físico experimentado, pueden ser frustrantes y confusos para la audiencia primaria de Feynman: los estudiantes.

Cabe resaltar que entre las 1 165 erratas que se corrigieron bajo mi auspicio son pocas las que considero como verdaderos errores en física. Un ejemplo está en el volumen II, página 5-10 que ahora dice “[...] si una cavidad está completamente encerrada por un conductor a tierra, ninguna distribución estática de cargas en el *interior* puede producir campos [eléctricos] en el *exterior*” (la precisión “a tierra” fue omitida en ediciones previas). Este error le fue señalado a Feynman por varios lectores, entre los que estaba Beulah Elizabeth Cox, estudiante en el College of William and Mary, quien confió en el pasaje erróneo de Feynman para un examen. En 1975 Feynman le escribió lo siguiente a la señorita Cox:† “Tu profesor tuvo razón en no darte puntos, pues tu respuesta es incorrecta, como lo demostró él usando la ley de Gauss. En la ciencia, se deben usar argumentos lógicos y elaborados con cuidado, no sólo creer en las autoridades. También leíste el libro correctamente y lo entendiste. Cometí un error, y el libro está mal. Tal vez estaba pensando en una esfera conductora a tierra, o bien en el hecho de que

† Páginas 288-289 de *Perfectly Reasonable Deviations from the Beaten Track, The Letters of Richard P. Feynman*, editado por Michelle Feynman (Basic Books, Nueva York, 2005).

mover las cargas en lugares distintos del interior no afecta las cosas en el exterior. No estoy seguro, pero metí la pata. Y tú también metiste la pata por creermelo”.

Cómo cobró vida esta Edición del nuevo milenio

Entre noviembre de 2005 y julio de 2006, 340 erratas fueron enviadas al sitio web de las *Lecciones de Feynman* www.feynmanlectures.info. Llama la atención que el grueso de éstas viniera de una sola persona: el doctor Rudolf Pfeiffer, que entonces era becario de posdoctorado en la Universidad de Viena, Austria. La editorial, Addison Wesley, arregló 80 erratas, pero se resistió a arreglar más debido al costo: los libros se imprimían mediante un proceso de foto offset y trabajaban con imágenes fotográficas de las páginas de los años sesenta. Corregir un error implicaba volver a formar la página entera, y para asegurar que no se colaran nuevos errores, la página debía volver a formarse por dos personas diferentes, para que luego la leyeran y corrigieran varias personas distintas: se trata de un proceso muy costoso cuando hay cientos de erratas en juego.

Gottlieb, Pfeiffer y Ralph Leighton estaban muy disgustados con esto, por lo que formularon un plan para facilitar el arreglo de todas las erratas, y también producir la versión electrónica y la aumentada de las *Lecciones de física de Feynman*. Me propusieron su plan a mí, como representante de Caltech, en 2007. Me emocioné, pero procedí con cautela. Después de ver más detalles, entre ellos un capítulo de muestra de la *Enhanced Electronic Version*, recomendé que Caltech cooperara con Gottlieb, Pfeiffer y Leighton para llevar a cabo su plan. El plan recibió la aprobación de tres catedráticos sucesivos de la División de Física, Matemáticas y Astronomía de Caltech: Tom Tombrello, Andrew Lange y Tom Soifer, y los detalles contractuales y legales complejos fueron resueltos por el abogado de Propiedad Intelectual de Caltech, Adam Cochran. Con la publicación de esta *Edición del nuevo milenio*, el plan se llevó a cabo con éxito, a pesar de su complejidad. Específicamente:

Pfeiffer y Gottlieb convirtieron a $\text{\LaTeX}$ los tres volúmenes de las *Lecciones* (y también más de 1000 ejercicios del curso de Feynman para incorporarlos a *Feynman's Tips on Physics*). Las figuras de las *Lecciones* se volvieron a trazar en formato electrónico moderno en India, bajo la supervisión del traductor al alemán de las *Lecciones*, Henning Heinze, para su uso en la edición alemana. Gottlieb y Pfeiffer intercambiaron el uso no exclusivo de sus ecuaciones en $\text{\LaTeX}$ en la edición alemana (publicada por Oldenbourg) por el uso no exclusivo de las figuras de Heinze en esta *Edición del nuevo milenio* inglesa. Pfeiffer y Gottlieb verificaron con cuidado todo el texto y ecuaciones en $\text{\LaTeX}$ y todas las figuras que se volvieron a trazar, e hicieron las correcciones necesarias. Nate Bode y yo, en representación de Caltech, hemos hecho controles aleatorios del texto, ecuaciones y figuras; y sorprendentemente, no hemos encontrado errores. Gottlieb y Pfeiffer son meticulosos y precisos en un grado increíble. Gottlieb y Pfeiffer concertaron que John Sullivan de la Biblioteca Huntington digitalizara las fotos de los pizarrones de Feynman entre 1962 y 1964, y que George Blood Audio digitalizara las grabaciones de las lecciones —con apoyo financiero y moral del profesor de Caltech Carver Mead, apoyo logístico de la archivista Shelley Erwin, y apoyo legal de Cochran—.

La problemática legal era seria: en los años sesenta, Caltech cedió los derechos a Addison Wesley para publicar la edición impresa, y en los años noventa, los derechos para distribuir el audio de las lecciones de Feynman y una variante de una edición electrónica. En la década del 2000, a través de una serie de adquisiciones de esas licencias, los derechos impresos se transfirieron al grupo editorial Pearson, mientras que los derechos electrónicos y de audio se transfirieron al grupo editorial Perseus. Con ayuda de Ike Williams, abogado especialista en lo editorial, Cochran logró unir todos estos derechos con Perseus (Basic Books), con lo que hizo posible esta *Edición del nuevo milenio*.

Agradecimientos

En nombre de Caltech, agradezco a las numerosas personas que han hecho posible esta *Edición del nuevo milenio*. En específico, quiero agradecer a las personas clave mencionadas líneas arriba: Ralph Leighton, Michael Gottlieb, Tom Tombrello, Michael Hartl, Rudolf Pfeiffer, Henning Heinze, Adam Cochran, Carver Mead, Nate Bode, Shelley Erwin, Andrew Lange, Tom Soifer, Ike Williams, y a las 50 personas que enviaron las erratas (quienes se enlistan en www.feynmanlectures.info). También agradezco a Michelle Feynman (hija de Richard Feynman) por su apoyo y consejo continuos, a Alan Rice por su asistencia tras bambalinas y asesoría en Caltech, a Stephan Puchegger y Calvin Jackson por su ayuda y orientación, a Pfeiffer por la conversión de las *Lecciones* a $\text{\LaTeX}$, y a Michael Figl, Manfred Smolik y Andreas Stangl por las discusiones sobre correcciones de erratas; así como al personal de Perseus/Basic Books, y (por las ediciones previas) al personal de Addison Wesley.

Kip S. Thorne
Profesor emérito Feynman de física teórica
Instituto de Tecnología de California

Octubre de 2010

Prefacio de Feynman



Éstas son las lecciones de física que di el año pasado y el antepasado a los estudiantes de los primeros años en el Instituto Tecnológico de California (Caltech). Por supuesto que éstas no son textuales; han sido editadas, a veces con gran extensión y a veces con muy poca. Las clases constituyen sólo parte del curso completo. El grupo total de 180 estudiantes se reunía dos veces por semana en un aula grande para atender a las explicaciones, y luego se dividía en grupos pequeños de 15 a 20 estudiantes en secciones de discusión y práctica bajo la guía de un ayudante de docencia. Había, además, una sesión de laboratorio por semana.

Con estas clases tratábamos de resolver un problema especial: mantener el interés de los estudiantes muy entusiastas y bastante despiertos que egresaban del bachillerato para entrar en el Caltech. Muchos habían oído hablar de lo interesante y estimulante que es la física: la teoría de la relatividad, la mecánica cuántica y otras ideas modernas. Pero al terminar los dos años del curso anterior al nuestro, muchos de ellos se sentían descorazonados porque realmente se les presentaban muy pocas ideas geniales, nuevas o interesantes. Se les hacía estudiar planos inclinados, electrostática y cuestiones por el estilo, y después de dos años era como para volverse tonto. Se trataba, pues, de ver si podíamos hacer un curso que salvara a los estudiantes más avanzados y animados manteniéndoles el entusiasmo.

Aunque mi intención no fue convertir las clases en un campo de estudio e investigación, preparé las lecciones para los más inteligentes de la clase a fin de asegurarme, si era posible, que aun los estudiantes más inteligentes no podrían abarcar completamente el contenido de cada lección; para ello introduje recomendaciones sobre la aplicación de las ideas y conceptos en diversas direcciones, las cuales se apartaban de la línea principal de ataque. Por esta razón, sin embargo, traté concienzudamente de hacer que todos los enunciados fueran lo más precisos posibles, de señalar en cada caso, dónde encajaban las ecuaciones y las ideas en el cuerpo de la física y cómo —cuando aprendieron más— se modificarían las cosas. Pensé, además, que para los estudiantes es importante indicar qué es lo que deben ser capaces —si son suficientemente inteligentes— de comprender por deducción de lo que se ha dicho antes y qué se está introduciendo como cosa nueva. Cuando se presentaban nuevas ideas, traté de deducirlas si eran deducibles o de explicar que *era* una nueva idea sin base alguna en lo que ya habían aprendido y que se suponía que no era demostrable, sino simplemente un agregado.

Al comenzar estas lecciones supuse que los estudiantes traían cierto conocimiento del bachillerato —como óptica geométrica, ideas simples sobre química, etc.—. Creí, también, que no había ninguna razón para dictar las lecciones en un orden determinado, en el sentido de que no debería mencionar algo hasta que no estuviera en condiciones de estudiarlo en detalle. Había abundante material que surgiría, pero sin discusión completa. Las discusiones completas se harían más tarde cuando hubiera mayor preparación para seguir las. Ejemplos de esto son el tratamiento de la inductancia y el de los niveles de energía, que primero se introducen en forma muy cualitativa y luego se desarrollan en forma más detallada.

Al mismo tiempo que me dirigía al estudiante más activo, quise cuidar del individuo para quien los adornos en demasía y las aplicaciones laterales son meramente intranquilizadores, cuando no se puede esperar, de ningún modo, que aprendan la mayor parte del material. Para ese estudiante traté que hubiera un núcleo central o columna vertebral de material que *pudiera* captar. Tenía la esperanza de que no se pondría nervioso aunque no entendiera todo el contenido de una lección. No esperaba que comprendiera todo sino los rasgos centrales y más directos. Naturalmente que se necesitaba cierta inteligencia de su parte para ver cuáles eran los teoremas e ideas centrales y cuáles los resultados y aplicaciones laterales más avanzados que sólo podría entender en años posteriores.

Había una dificultad seria para dar estas lecciones: por la forma en que se daba el curso, no había una retroacción del estudiante hacia el profesor que indicara cómo estaban yendo las lecciones. Ésta es una dificultad muy seria que me impide saber con certeza hasta qué punto, en realidad, fueron muy provechosas mis clases. Todo era esencialmente un experimento. Y si lo hiciera nuevamente no lo haría en la misma forma —¡espero que *no* tenga que hacerlo de nuevo!— Creo, no obstante, que en lo que respecta a la física las cosas anduvieron muy satisfactoriamente el primer año.

El segundo año no quedé tan satisfecho. En la primera parte del curso, que trataba de la electricidad y el magnetismo, no pude encontrar ningún modo realmente único o diferente de hacerlo —ninguna manera que fuera especialmente más estimulante que la forma habitual de presentarlo—. Por lo tanto, no creo que hice mucho en las clases sobre electricidad y magnetismo. Al final del segundo año había pensado originalmente continuar dando, después de la electricidad y el magnetismo, algunas clases sobre las propiedades de los materiales, no obstante con el interés de explicar los modos fundamentales, las soluciones de la ecuación de difusión, los sistemas vibrantes, las funciones ortogonales... desarrollando, así, las primeras etapas de lo que usualmente se denomina “métodos matemáticos de la física”. Pensando en retrospectiva, creo que si lo hiciera de nuevo volvería a esa idea original. Pero como no se había contemplado que yo daría estas clases nuevamente, se sugirió que sería conveniente tratar de dar una introducción a la mecánica cuántica —que es lo que ustedes encontrarán en el tercer volumen—.

Queda perfectamente claro que los estudiantes que sigan estudiando física deberán esperar hasta el tercer año para estudiar mecánica cuántica. Por otra parte, se esgrimió el argumento de que muchos de los estudiantes de nuestro curso estudian física como base para su especialización en otros campos. Y la forma habitual de tratar la mecánica cuántica hace que el tema sea casi inalcanzable para la gran mayoría de los estudiantes ya que necesitan mucho tiempo para aprenderlo. No obstante, en sus aplicaciones concretas —especialmente en sus aplicaciones más complejas, como en la ingeniería eléctrica y en la química— realmente no se usa la maquinaria completa del tratamiento con ecuaciones diferenciales. Por ello, traté de describir los principios de la mecánica cuántica de un modo que no exigiera un conocimiento básico de la matemática de las ecuaciones diferenciales. Creo que aun para un físico es muy interesante presentar la mecánica cuántica de esta manera inversa —por varias razones que se pueden ver en las lecciones—. Sin embargo, creo que el experimento en lo correspondiente a la mecánica cuántica no tuvo éxito completo —en gran parte debido a que no tuve tiempo al final (por ejemplo, debería haber tenido tres o cuatro lecciones más para tratar detenidamente temas como las bandas de energía y la dependencia

espacial de las amplitudes)—. Además, nunca había presentado antes el tema de este modo, por lo que la falta de retroacción fue particularmente seria. Ahora creo que se debe dar la mecánica cuántica más tarde. A lo mejor algún día tenga la oportunidad de hacerlo de nuevo. Entonces lo haré mejor.

La razón de que no haya lecciones sobre cómo resolver problemas se debe a que había secciones de discusión y práctica. Aunque en el primer año incluí tres lecciones sobre cómo resolver problemas, en este curso no pude hacerlo. También había una lección sobre guía inercial que debe estar ciertamente después de la clase sobre sistemas que rotan, pero que desafortunadamente se omitió. Las lecciones quinta y sexta fueron dadas por Matthew Sands porque yo estaba fuera de la ciudad.

La pregunta es, por supuesto, hasta qué punto este experimento ha tenido éxito. Mi punto de vista —que, sin embargo, no lo comparten la mayoría de los que trabajaron con los estudiantes— es pesimista. No creo que me haya ido muy bien con los estudiantes. Cuando recuerdo el modo en que éstos manipulaban los problemas en los exámenes, pienso que el sistema es un fracaso. Por supuesto, mis amigos me indican que hubo una o dos docenas de estudiantes que —muy sorprendentemente— comprendieron casi todo el contenido en todas las lecciones y que fueron muy activos trabajando con el material y preocupándose con animación e interés por muchos tópicos. Estos individuos tienen actualmente, creo, una base de primera línea en física —y son, después de todo, aquellos a quienes quería llegar—. Pero entonces “El poder de la instrucción es, en general, poco eficaz, excepto en las felices disposiciones en que es casi superfluo” (Gibbons).

De todos modos, no quería dejar ningún estudiante completamente atrasado, como quizá lo hice. Creo que un modo más efectivo de ayudar al estudiante sería mediante la adición de tiempo y esfuerzo en el desarrollo de un conjunto de problemas que aclare algunas de las ideas contenidas en las lecciones. Los problemas dan la oportunidad de aumentar la comprensión del material expuesto haciéndolo más real, estructurado y accesible para el proceso de fijación.

Pienso, sin embargo, que la solución a este problema educativo no es otra que darse cuenta que la enseñanza sólo puede realizarse cuando hay una relación individual directa entre un estudiante y un buen profesor, situación en la cual el estudiante discute las ideas, piensa en las cosas y habla sobre ellas. Es imposible aprender simplemente asistiendo a una clase, o simplemente resolviendo los problemas asignados. Pero en los actuales momentos tenemos tantos estudiantes a quienes enseñar que debemos encontrar un sustituto de lo ideal. Quizá mis lecciones ejerzan alguna contribución. Quizá en algún lugar pequeño donde sea posible una relación individual entre profesores y estudiantes, éstos obtengan alguna inspiración o algunos conceptos de estas lecciones. Quizá entonces, también el proceso de fijar el material sea más alegre y placentero para ellos y dé origen al desarrollo de algunas ideas.

RICHARD P. FEYNMAN

Junio, 1963

Introducción

Este libro está basado en el curso de lecciones introductorias de física que dio el profesor R. P. Feynman en el Instituto de Tecnología de California (Caltech) durante el año académico 1961-1962; abarca el primero de los dos años del curso introductorio que toman todos los estudiantes de primer y segundo años en Caltech, y a éste siguió en 1962-1963 una serie similar que abarcó el segundo año. Las lecciones constituyen una parte muy importante en la revisión del curso introductorio, la cual se llevó a cabo a lo largo de un periodo de cuatro años.

La necesidad de una revisión básica surgió tanto del rápido desarrollo de la física en las décadas recientes como del hecho de que los estudiantes recién llegados han mostrado un aumento constante en su habilidad matemática a partir de las mejoras en el contenido de las clases de matemáticas del bachillerato. Nuestra intención fue aprovechar este mayor bagaje matemático y al mismo tiempo introducir suficientes temas modernos para que el curso fuera exigente, interesante y más representativo de la física de hoy en día.

Para generar una variedad de ideas sobre qué material incluir y cómo presentarlo, se alentó a un número sustancial de profesores de física a que ofrecieran sus ideas en forma de planes temáticos para un curso revisado. Varios de éstos fueron presentados y discutidos de manera crítica y profunda. Casi de inmediato se llegó al acuerdo de que una revisión básica del curso no se podría lograr adoptando un libro de texto diferente, ni siquiera escribiendo uno *ab initio*, sino que el nuevo curso debía centrarse en un conjunto de lecciones presentadas a ritmo de dos o tres por semana; el material textual apropiado se produciría luego como una operación secundaria conforme avanzaba el curso, y también se organizarían experimentos de laboratorio adecuados para el material de las lecciones. En consecuencia, se fijó un plan del curso a grandes rasgos, pero se reconoció que era tentativo y estaba incompleto y sujeto a modificaciones considerables por parte de quien fuera a tener la responsabilidad de efectivamente preparar las lecciones.

En cuanto al mecanismo mediante el cual sería finalmente puesto en práctica el curso, se consideraron varios planes. Éstos se parecían mucho entre sí, incluían un esfuerzo cooperativo por N número de miembros del personal que compartirían la carga total de trabajo de manera simétrica y equitativa: cada uno tomaría la responsabilidad de $1/N$ del material, daría las lecciones y escribiría el material textual para su parte. Sin embargo, la falta de disponibilidad de personal suficiente y la dificultad de mantener un punto de vista uniforme ante las diferencias de personalidad y filosofía entre los participantes individuales hacían que estos planes fueran irrealizables.

Darse cuenta de que verdaderamente poseíamos los medios para crear no sólo un curso nuevo y diferente de física, sino quizá uno único, resultó ser una inspiración dichosa para el profesor Sands. Él sugirió que el profesor R. P. Feynman preparara y diera las lecciones, y que éstas fueran grabadas. Una vez transcritas y editadas, se convertirían en el libro de texto para el nuevo curso. En esencia, éste fue el plan que se adoptó.

Se esperaba que la edición necesaria fuera menor, pues consistiría principalmente en proporcionar figuras y revisar la puntuación y la gramática; la harían uno o dos estudiantes de posgrado a tiempo parcial. Desafortunadamente, esta expectativa fue pasajera. En realidad, fue una operación editorial mayor poder transformar la transcripción literal a una forma legible, aun sin la reorganización o revisión del tema tratado que a veces se requería. Además, no era trabajo para un editor técnico o para un estudiante de posgrado, sino uno que requería la atención minuciosa de un físico profesional ¡de 10 a 20 horas por lección!

La dificultad de la tarea editorial, junto con la necesidad de poner el material en manos de los estudiantes tan pronto como fuera posible, puso un límite estricto en qué tanto se podía “pulir” el material en realidad, y en consecuencia nos vimos forzados a apostar por un producto preliminar pero técnicamente correcto que pudiera usarse de inmediato, en lugar de uno que pudiera considerarse final o terminado. Debido a una necesidad urgente de más ejemplares para nuestros alumnos, y a un interés alentador por parte de profesores y estudiantes de varias instituciones distintas, decidimos publicar el material en su forma preliminar en lugar de esperar una revisión ulterior más profunda que quizá nunca llegaría. No nos hacemos ilusiones en cuanto a lo completo o fluido del material, ni en cuanto a su organización lógica; de hecho, planeamos varias modificaciones menores en el futuro inmediato, y esperamos que no se vuelva estático en su forma o contenido.

Además de las lecciones, que constituyen una parte de importancia central para el curso, era necesario también proporcionar ejercicios adecuados para acrecentar la experiencia y habilidad de los estudiantes, así como experimentos oportunos para facilitar un contacto de primera mano con el material de las lecciones en el laboratorio. Ninguno de estos aspectos se encuentra en un estado tan avanzado como el material de las lecciones, pero se ha hecho un progreso considerable. Algunos ejercicios fueron elaborados conforme progresaban las lecciones y luego expandidos y amplificados para su uso al año siguiente. Sin embargo, puesto que todavía no estamos convencidos de que los ejercicios proporcionen suficiente variedad y profundidad de aplicación de los materiales de las lecciones para dar al alumno plena conciencia del tremendo poder que está a su disposición, los ejercicios se publicaron por separado en un formato menos permanente para así alentar una revisión frecuente.

El profesor H. V. Neher ha ideado varios novedosos experimentos para el nuevo curso. Entre ellos se encuentran varios que utilizan la fricción extremadamente baja que presenta un cojinete de gas: un nuevo canal lineal de aire con el que es posible hacer mediciones cuantitativas de movimiento unidimensional, impactos y movimiento armónico, y un trompo de color de Maxwell soportado e impulsado por aire con el que se pueden estudiar el movimiento de rotación acelerado y la precesión y nutación giroscópicas. Se espera que el desarrollo de nuevos experimentos de laboratorio dure un tiempo considerable.

El programa de revisión se llevó a cabo bajo la dirección de los profesores R. B. Leighton, H. V. Neher y M. Sands. En el programa participaron oficialmente los profesores R. P. Feynman, G. Neugebauer, R. M. Sutton, H. P. Stabler, F. Strong y R. Vogt, de la división de física, matemáticas y astronomía, y los profesores T. Caughey, M. Plesset y C. H. Wilts de la división de ciencias de la ingeniería. Reconocemos y agradecemos la valiosa asistencia de todos aquellos que contribuyeron al programa de revisión. Estamos en deuda en especial con la Fundación Ford, sin cuyo apoyo financiero este programa no podría haberse llevado a cabo.

ROBERT B. LEIGHTON

Julio, 1963

Índice general

CAPÍTULO 1. COMPORTAMIENTO CUÁNTICO

1-1 Mecánica atómica	1-1
1-2 Un experimento con balas	1-2
1-3 Un experimento con ondas	1-3
1-4 Un experimento con electrones	1-4
1-5 Interferencia de ondas-electrones	1-6
1-6 Espiando a los electrones	1-7
1-7 Los primeros principios de la mecánica cuántica	1-10
1-8 El principio de incertidumbre	1-11

CAPÍTULO 2. RELACIÓN ENTRE LOS PUNTOS DE VISTA DE ONDAS Y DE PARTÍCULAS

2-1 Amplitudes de onda de probabilidad	2-1
2-2 Medición de la posición y el momentum	2-2
2-3 Difracción de cristales	2-5
2-4 El tamaño de un átomo	2-6
2-5 Niveles de energía	2-7
2-6 Implicaciones filosóficas	2-8

CAPÍTULO 3. AMPLITUDES DE PROBABILIDAD

3-1 Leyes para combinar amplitudes	3-1
3-2 El patrón de interferencia de dos rejillas	3-5
3-3 Dispersión en un cristal	3-8
3-4 Partículas idénticas	3-10

CAPÍTULO 4. PARTÍCULAS IDÉNTICAS

4-1 Bosones y fermiones	4-1
4-2 Estados con dos bosones	4-3
4-3 Estados con n bosones	4-6
4-4 Emisión y absorción de fotones	4-8
4-5 El espectro de cuerpo negro	4-9
4-6 El helio líquido	4-12
4-7 El principio de exclusión	4-13

CAPÍTULO 5. ESPÍN UNO

5-1 Filtrando átomos con un aparato de Stern-Gerlach	5-1
5-2 Experimentos con átomos filtrados	5-5
5-3 Filtros de Stern-Gerlach en serie	5-7
5-4 Estados base	5-8
5-5 Amplitudes que interfieren	5-10

5-6 La maquinaria de la mecánica cuántica	5-13
5-7 Transformando a una base diferente	5-15
5-8 Otras situaciones	5-17

CAPÍTULO 6. ESPÍN UN MEDIO

6-1 Transformando amplitudes	6-1
6-2 Transformando a un sistema de coordenadas rotado	6-3
6-3 Rotaciones alrededor del eje z	6-6
6-4 Rotaciones de 180° y de 90° alrededor de y	6-9
6-5 Rotaciones alrededor de x	6-12
6-6 Rotaciones arbitrarias	6-13

CAPÍTULO 7. DEPENDENCIA TEMPORAL DE LAS AMPLITUDES

7-1 Átomos en reposo, estados estacionarios	7-1
7-2 Movimiento uniforme	7-3
7-3 Energía potencial, conservación de la energía	7-6
7-4 Fuerzas: el límite clásico	7-9
7-5 La “precesión” de una partícula con espín un medio	7-11

CAPÍTULO 8. LA MATRIZ HAMILTONIANA

8-1 Amplitudes y vectores	8-1
8-2 Descomposición de vectores de estado	8-3
8-3 ¿Cuáles son los estados base del mundo?	8-5
8-4 ¿Cómo cambian los estados en el tiempo?	8-8
8-5 La matriz hamiltoniana	8-10
8-6 La molécula de amoníaco	8-11

CAPÍTULO 9. EL MÁSER DE AMONIACO

9-1 Los estados de una molécula de amoníaco	9-1
9-2 La molécula en un campo eléctrico estático	9-5
9-3 Transiciones en un campo dependiente del tiempo	9-9
9-4 Transiciones en resonancia	9-12
9-5 Transiciones fuera de resonancia	9-14
9-6 Absorción de luz	9-15

CAPÍTULO 10. OTROS SISTEMAS DE DOS ESTADOS

10-1 El ion del hidrógeno molecular	10-1
10-2 Fuerzas nucleares	10-6
10-3 La molécula de hidrógeno	10-8
10-4 La molécula de benceno	10-11
10-5 Colorantes	10-12

10-6 El hamiltoniano de una partícula con espín un medio en un campo magnético	10-13
10-7 El electrón con espín en un campo magnético	10-16

CAPÍTULO 11. MÁS SISTEMAS DE DOS ESTADOS

11-1 Las matrices de espín de Pauli	11-1
11-2 Las matrices de espín como operadores	11-5
11-3 La solución de las ecuaciones de dos estados	11-8
11-4 Los estados de polarización del fotón	11-9
11-5 El mesón K neutro	11-13
11-6 Generalización a sistemas de N estados	11-21

CAPÍTULO 12. LA ESTRUCTURA HIPERFINA EN EL HIDRÓGENO

12-1 Estados base para un sistema de dos partículas con espín un medio	12-1
12-2 El hamiltoniano para el estado fundamental del hidrógeno	12-3
12-3 Los niveles de energía	12-7
12-4 La separación de Zeeman	12-9
12-5 Los estados en un campo magnético	12-13
12-6 La matriz de proyección para espín uno	12-15

CAPÍTULO 13. PROPAGACIÓN EN UNA RED CRISTALINA

13-1 Estados para un electrón en una red unidimensional	13-1
13-2 Estados de energía definida	13-3
13-3 Estados dependientes del tiempo	13-7
13-4 Un electrón en una red tridimensional	13-8
13-5 Otros estados en una red	13-9
13-6 Dispersión por imperfecciones en la red	13-10
13-7 Atrapar con una imperfección en la red	13-12
13-8 Amplitudes de dispersión y estados enlazados	13-13

CAPÍTULO 14. SEMICONDUCTORES

14-1 Electrones y huecos en semiconductores	14-1
14-2 Semiconductores impuros	14-4
14-3 El efecto Hall	14-7
14-4 Junturas de semiconductores	14-8
14-5 Rectificación en una juntura de semiconductores	14-11
14-6 El transistor	14-12

CAPÍTULO 15. LA APROXIMACIÓN DE PARTÍCULAS INDEPENDIENTES

15-1 Ondas de espín	15-1
15-2 Dos ondas de espín	15-4
15-3 Partículas independientes	15-6
15-4 La molécula de benceno	15-7
15-5 Más química orgánica	15-11
15-6 Otros usos para la aproximación	15-13

CAPÍTULO 16. LA DEPENDENCIA DE LAS AMPLITUDES CON LA POSICIÓN

16-1 Amplitudes sobre una línea	16-1
16-2 La función de onda	16-5
16-3 Estados de momentum definido	16-7
16-4 Normalización de los estados en x	16-10
16-5 La ecuación de Schrödinger	16-12
16-6 Niveles de energía cuantizados	16-15

CAPÍTULO 17. SIMETRÍA Y LEYES DE CONSERVACIÓN

17-1 Simetría	17-1
17-2 Simetría y conservación	17-3
17-3 Las leyes de conservación	17-7
17-4 Luz polarizada	17-10
17-5 La desintegración de la Λ^0	17-12
17-6 Resumen de las matrices de rotación	17-16

CAPÍTULO 18. MOMENTUM ANGULAR

18-1 Radiación dipolar eléctrica	18-1
18-2 Dispersión de la luz	18-3
18-3 La aniquilación del positronio	18-5
18-4 Matriz de rotación para un espín cualquiera	18-10
18-5 Midiendo un espín nuclear	18-13
18-6 Composición de momentum angular	18-15
18-7 Nota adicional 1: obtención de la matriz de rotación	18-20
18-8 Nota adicional 2: conservación de la paridad en la emisión de fotones	18-23

CAPÍTULO 19. EL ÁTOMO DE HIDRÓGENO Y LA TABLA PERIÓDICA

19-1 La ecuación de Schrödinger para el átomo de hidrógeno	19-1
19-2 Soluciones con simetría esférica	19-2
19-3 Estados con dependencia angular	19-6
19-4 Solución general para el hidrógeno	19-9
19-5 Funciones de onda del hidrógeno	19-13
19-6 La tabla periódica	19-14

CAPÍTULO 20. OPERADORES

20-1 Operaciones y operadores	20-1
20-2 Valores medios de la energía	20-3
20-3 Valor medio de la energía de un átomo	20-6
20-4 El operador posición	20-8
20-5 El operador momentum	20-9
20-6 Momentum angular	20-14
20-7 Variación temporal de valores medios	20-16

CAPÍTULO 21. LA ECUACIÓN DE SCHRÖDINGER
EN EL CONTEXTO CLÁSICO: UN SEMINARIO
SOBRE SUPERCONDUCTIVIDAD

21-1 La ecuación de Schrödinger en un campo magnético	21-1
21-2 Ecuación de continuidad para probabilidades . .	21-3
21-3 Dos clases de momentum	21-5
21-4 El significado de la función de onda	21-6
21-5 Superconductividad	21-7
21-6 El efecto Meissner	21-8

21-7 Cuantificación del flujo	21-10
21-8 Dinámica de la superconductividad	21-12
21-9 Juntura de Josephson	21-14

Epílogo de Feynman

Apéndice

Índice analítico

Índice de nombres

Lista de símbolos

Comportamiento cuántico

1-1 Mecánica atómica

La “mecánica cuántica” es la descripción del comportamiento de la materia y de la luz en todos sus detalles y particularmente de los acontecimientos en escala atómica. Las cosas en una escala muy pequeña no se comportan como nada que ustedes conozcan directamente. No se comportan como ondas, no se comportan como partículas, no se comportan como nubes, o como bolas de billar, o como pesas en resortes, o como algo que hayan visto alguna vez.

Newton pensaba que la luz estaba compuesta por partículas, luego se descubrió que ésta se comporta como una onda. Más tarde, sin embargo (a principios del siglo XX), se descubrió que la luz algunas veces sí se comportaba como una partícula. Históricamente, el electrón, por ejemplo, se pensaba que se comportaba como una partícula, y luego se descubrió que, en muchos aspectos, se comportaba como una onda. En realidad, no se comporta como ninguna de ellas. Ahora nos hemos rendido. Decimos que: “No es como *ninguna de las dos*”.

Sin embargo, tenemos algo de suerte —los electrones se comportan exactamente como la luz—. El comportamiento cuántico de objetos atómicos (electrones, protones, neutrones, fotones, entre otros) es el mismo para todos ellos, todos son “ondas-partículas” o como quieran llamarlas. De este modo, todo lo que aprendamos sobre las propiedades de los electrones (los cuales usaremos para nuestros ejemplos) se puede aplicar a todas las “partículas”, incluyendo a los fotones de luz.

La acumulación gradual de información acerca del comportamiento atómico y en pequeña escala durante el primer cuarto del siglo XX, dio algunas indicaciones de cómo se comportan en realidad las cosas pequeñas y produjo una confusión creciente que fue resuelta finalmente en 1926 y 1927 por Schrödinger, Heisenberg y Born. Ellos obtuvieron una descripción consistente del comportamiento de la materia en una escala pequeña. Nosotros retomaremos los principales rasgos de esa descripción en este capítulo.

Debido a que el comportamiento atómico es tan diferente a nuestra experiencia ordinaria, es muy difícil acostumbrarse a éste, es peculiar y misterioso para todos —tanto para el físico novato como para el experimentado—. Incluso los expertos no lo comprenden en la forma que ellos quisieran, y es perfectamente razonable que ocurra así, porque la experiencia directa y la intuición humana se aplican sólo a los objetos grandes. Sabemos cómo se comportan los objetos grandes, pero los objetos en escala pequeña no se comportan igual. Así que tenemos que aprender de ellos en una forma un tanto abstracta o imaginaria, y sin relacionarlos con nuestra experiencia directa.

En este capítulo abordaremos inmediatamente el elemento básico de este comportamiento tan misterioso en su forma más extraña. Nosotros elegimos examinar un fenómeno que es imposible, *absolutamente* imposible, de explicar de manera clásica, y que encierra en sí el corazón de la mecánica cuántica. En realidad, contiene el *único* misterio. No podemos desaparecer el misterio “explicando” cómo funciona. Sólo les *diremos* cómo funciona. Al decirles cómo funciona, les hablamos acerca de las peculiaridades básicas de toda la mecánica cuántica.

1-1 Mecánica atómica

1-2 Un experimento con balas

1-3 Un experimento con ondas

1-4 Un experimento con electrones

1-5 Interferencia de ondas-electrones

1-6 Espiando a los electrones

1-7 Los primeros principios de la mecánica cuántica

1-8 El principio de incertidumbre

Nota: Este capítulo es casi lo mismo que el capítulo 37 del volumen I.

1-2 Un experimento con balas

A fin de entender el comportamiento cuántico de los electrones, compararemos y contrastaremos su comportamiento, en un arreglo experimental determinado, con el comportamiento más familiar de partículas como balas, y con el comportamiento de ondas como las ondas en el agua. Consideraremos primero el comportamiento de las balas en el arreglo experimental que se muestra a manera de diagrama en la figura 1-1. Tenemos una ametralladora que dispara un chorro de balas. No es una ametralladora muy buena, pues dispersa los proyectiles (de manera aleatoria) en un ángulo bastante grande, como se indica en la figura. Delante de la ametralladora tenemos una pared (hecha con una placa blindada) que posee dos agujeros, lo suficientemente grandes para que pase una bala. Después de la pared hay una barrera (digamos una pared gruesa de madera) que “absorberá” las balas cuando éstas la golpeen. Frente a la pared tenemos un objeto al que llamaremos “detector” de balas. Puede ser una caja que contenga arena. Cualquier bala que entre en el detector se detendrá y se irá acumulando. Cuando lo deseemos, podemos vaciar la caja y contar el número de balas acumuladas. El detector se puede mover hacia adelante y hacia atrás (en la dirección que llamaremos dirección x). Con este aparato podemos encontrar experimentalmente la respuesta a la siguiente pregunta: “¿Cuál es la probabilidad para que una bala que pasa a través de los agujeros en la pared llegue a la barrera con distancia x del centro?” Primero, deben darse cuenta de que tenemos que hablar de probabilidad, ya que no podemos decir exactamente a dónde irá una bala determinada. Una bala que golpee uno de los agujeros puede rebotar en el borde y terminar en cualquier lugar. Por “probabilidad”, entendemos la posibilidad de que la bala llegue al detector, la cual podemos medir contando el número de balas que llegan al detector en un cierto tiempo, y dividiendo este número entre el número *total* que inciden sobre la barrera durante el mismo tiempo. O bien, si suponemos que la ametralladora siempre dispara con la misma rapidez mientras se mide, la probabilidad que queremos es proporcional al número de balas que llega al detector en algún intervalo estándar de tiempo.

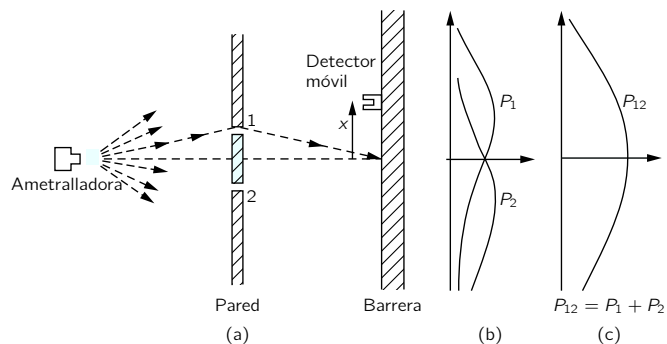


Figura 1-1. Experimento de interferencia con balas.

Para los fines que nos proponemos, nos gustaría imaginar un experimento idealizado en el cual las balas no fuesen balas reales, sino balas *indestructibles* —éstas no se pueden partir por la mitad—. En nuestro experimento encontramos que las balas siempre llegan en paquetes, y cuando encontramos algo en el detector, siempre es una bala completa. Si la rapidez a la cual la ametralladora dispara es muy pequeña, encontramos que en un momento dado ya sea que ninguna, o una y sólo una —exactamente una— bala llega a la barrera. Además, el tamaño del paquete no depende de la rapidez de disparo de la ametralladora. Diremos: “Las balas *siempre* llegan en paquetes idénticos”. Lo que medimos con nuestro detector es la probabilidad de llegada de un paquete. Y medimos dicha probabilidad en función de x . El resultado de tales mediciones con este aparato (aún no hemos realizado el experimento, así que en realidad estamos imaginando el resultado) está representado en la gráfica de la parte (c) de la figura 1-1. En la gráfica representamos a la probabilidad en la derecha y a x verticalmente, de tal forma que la escala de x se ajuste al diagrama del aparato. Llamamos a la

probabilidad P_{12} , porque las balas pueden haber llegado a través del agujero 1 o a través del agujero 2. No se sorprendan de que P_{12} sea más grande y cercano al centro de la gráfica, y que disminuya si x aumenta. Pueden preguntar, sin embargo, por qué P_{12} alcanza su valor máximo en $x = 0$. Podemos entender este hecho si repetimos nuestro experimento después de haber tapado el agujero 2, y una vez más tapando ahora el agujero 1. Cuando tapamos el agujero 2, las balas sólo pueden pasar a través del agujero 1, y obtenemos la curva marcada como P_1 en la parte (b) de la figura. Como podrían esperar, el máximo de P_1 ocurre en el valor de x que está en línea recta con la ametralladora y el agujero 1. Cuando tapamos el agujero 1, obtenemos la curva simétrica P_2 dibujada en la figura. P_2 es la distribución de probabilidad para las balas que atraviesan el agujero 2. Al comparar las partes (b) y (c) de la figura 1-1, encontramos el siguiente resultado importante:

$$P_{12} = P_1 + P_2. \quad (1.1)$$

Las probabilidades sólo se suman. El efecto que se obtiene con los dos agujeros abiertos es la suma de los efectos obtenidos con cada agujero solo. Llamaremos a este resultado una observación “sin interferencia”, por una razón que verán más adelante. Hasta aquí con las balas. Éstas llegan en paquetes, y su probabilidad de llegada no presenta interferencia.

1-3 Un experimento con ondas

Ahora deseamos considerar un experimento con ondas de agua. El aparato se muestra de manera esquemática en la figura 1-2. Tenemos un recipiente poco profundo con agua. Un pequeño objeto, denominado “fuente de ondas”, es agitado mediante un motor y forma ondas circulares. A la derecha de la fuente tenemos de nuevo una pared con dos agujeros, y después una segunda pared, la cual, para simplificar las cosas, diremos que es un “amortiguador”, de tal modo que no se reflejan las ondas que llegan a él. Esto se puede hacer construyendo una “playa” gradual de arena. Delante de la playa colocamos un detector que podemos mover hacia un lado o hacia otro en la dirección x , como en el caso anterior. El detector es ahora un aparato que mide la “intensidad” del movimiento de las ondas. Pueden imaginar que es un mecanismo que mide la altura de las ondas, pero cuya escala está calibrada en proporción con el *cuadrado* de la altura en cuestión, de tal manera que la lectura es proporcional a la intensidad de la onda. Nuestro detector lee, entonces, en proporción a la *energía* transportada por la onda —o mejor, la rapidez con que la energía es transportada al detector—.

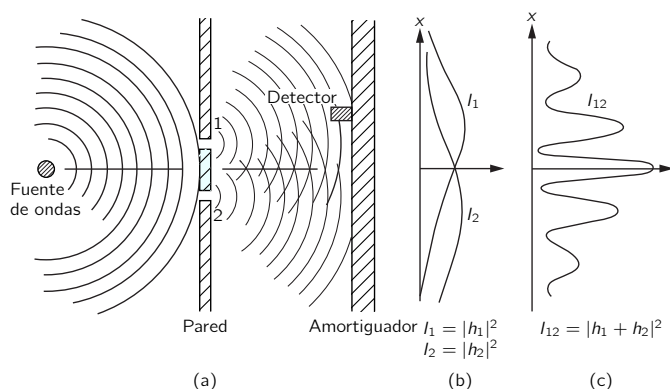


Figura 1-2. Experimento de interferencia con ondas de agua.

Con nuestro aparato de ondas, lo primero que debemos notar es que la intensidad puede tener *cualquier tamaño*. Si la fuente se mueve sólo un poco, entonces el movimiento de la onda en el detector es muy pequeño. Cuando hay más movimiento en la fuente, hay más intensidad en el detector. La intensidad de la onda puede tener cualquier valor. *No* podríamos decir que existe algún “paquete” de intensidad de onda.

Ahora midamos la intensidad de onda para diferentes valores x (conservando la fuente de ondas operando de la misma forma). Obtenemos la interesante curva marcada como I_{12} en la parte (c) de la figura.

Ya vimos cómo se pueden producir tales patrones cuando estudiamos la interferencia de ondas eléctricas en el volumen I. En este caso hemos observado que la onda original se difracta en los agujeros, y que nuevas ondas circulares surgen en cada agujero. Si tapamos un agujero a la vez y medimos la distribución de intensidad en el amortiguador, encontramos las curvas de intensidad más simples de la parte (b) de la figura. Así, I_1 es la intensidad de la onda del agujero 1 (la cual encontramos midiendo cuando tapamos el agujero 2) e I_2 es la intensidad de la onda del agujero 2 (obtenida cuando tapamos el agujero 1).

La intensidad I_{12} observada cuando ambos agujeros están abiertos ciertamente *no* es la suma de I_1 e I_2 . Decimos que hay “interferencia” de las dos ondas. En algunos lugares (donde la curva I_{12} tiene sus máximos) las ondas están “en fase” y los picos de las ondas se suman para dar una mayor amplitud y, por tanto, una mayor intensidad. Decimos que las dos ondas están “interfiriendo de manera constructiva” en esos lugares. Existirá dicha interferencia constructiva donde la distancia del detector a un agujero sea un múltiplo entero de las longitudes de onda mayor (o menor) que la distancia del detector al otro agujero.

En los lugares donde las dos ondas llegan al detector con una diferencia de fase de π (donde están “fuera de fase”), el movimiento de la onda resultante en el detector será la diferencia de las dos amplitudes. Las ondas “interfieren de manera destructiva” y obtenemos un valor bajo para la intensidad de onda. Esperamos esos valores bajos donde la distancia entre el agujero 1 y el detector difiera de la distancia entre el agujero 2 y el detector en un múltiplo impar de las semilongitudes de onda. Los valores mínimos de I_{12} en la figura 1-2 corresponden a los lugares donde las dos ondas interfieren de manera destructiva.

Ustedes recordarán que la relación cuantitativa entre I_1 , I_2 e I_{12} se puede expresar en la forma siguiente: la altura instantánea de la onda de agua en el detector para la onda del agujero 1 se puede escribir como (la parte real de) $h_1 e^{i\omega t}$, donde la “amplitud” h_1 es, en general, un número complejo. La intensidad es proporcional al cuadrado de la altura media o, cuando usamos números complejos, al cuadrado de la norma $|h_1|^2$. Del mismo modo, para el agujero 2, la altura es $h_2 e^{i\omega t}$ y la intensidad es proporcional a $|h_2|^2$. Cuando los dos agujeros están abiertos, las alturas de las ondas se suman para dar la altura $(h_1 + h_2)e^{i\omega t}$ y la intensidad $|h_1 + h_2|^2$. Omitimos la constante de proporcionalidad para nuestros propósitos actuales, las relaciones apropiadas para *ondas que interfieren* son:

$$I_1 = |h_1|^2, \quad I_2 = |h_2|^2, \quad I_{12} = |h_1 + h_2|^2. \quad (1.2)$$

Notarán que el resultado es bastante diferente al que obtuvimos con las balas (ecuación 1.1). Si desarrollamos $|h_1 + h_2|^2$ vemos que:

$$|h_1 + h_2|^2 = |h_1|^2 + |h_2|^2 + 2|h_1||h_2|\cos\delta, \quad (1.3)$$

donde δ es la diferencia de fase entre h_1 y h_2 . En términos de las intensidades podemos escribir

$$I_{12} = I_1 + I_2 + 2\sqrt{I_1 I_2} \cos\delta. \quad (1.4)$$

El último término en (1.4) es el “término de interferencia”. Hasta aquí con las ondas de agua. La intensidad puede tener cualquier valor, y muestran interferencia.

1-4 Un experimento con electrones

Ahora imaginemos un experimento análogo con electrones. Éste se muestra en forma de diagrama en la figura 1-3. Hacemos un cañón de electrones, el cual consiste en un alambre de tungsteno calentado mediante una corriente eléctrica y rodeado por una caja de metal que tiene un agujero en ella. Si el alambre tiene un potencial eléctrico negativo respecto a la caja, los electrones emitidos por el

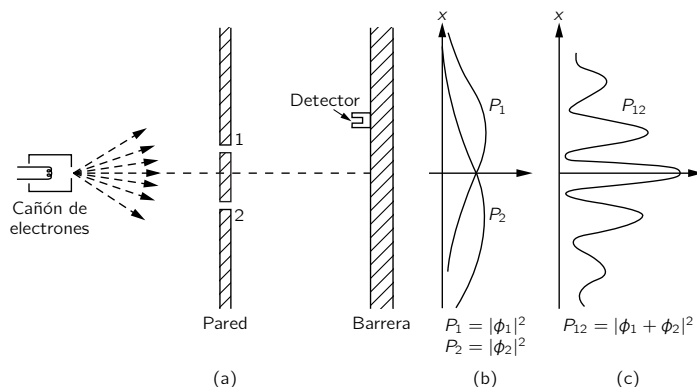


Figura 1-3. Experimento de interferencia con electrones.

alambre se acelerarán hacia las paredes y algunos atravesarán el agujero. Todos los electrones que salgan del cañón tendrán (aproximadamente) la misma energía. Delante del cañón hay de nuevo una pared (una delgada placa de metal) con dos agujeros. Después de la pared hay otra placa que funciona como “barrera”. En frente de la barrera colocamos un detector móvil. El detector puede ser un contador Geiger o, quizá mejor, un multiplicador de electrones, que se encuentra conectado a un altavoz.

Debemos ya advertir que no deben intentar montar este experimento (como pudieron haber hecho con los dos descritos anteriormente). Este experimento nunca se ha realizado de esta manera. El problema es que el aparato tendría que construirse en una escala imposiblemente pequeña para que pudiera mostrar los efectos que nos interesan. Estamos haciendo un “experimento mental”, que hemos elegido porque es fácil de imaginar. Sabemos los resultados que *se obtendrían* porque *hay* muchos experimentos que se han llevado a cabo, en los cuales la escala y las proporciones se han escogido para mostrar los efectos que describiremos.

Lo primero que notamos con nuestro experimento con electrones es que escuchamos “clics” bien definidos provenientes del detector (es decir, del altavoz). Y todos los “clics” son iguales. *No* hay “medios clics”.

También notamos que los “clics” ocurren muy al azar. Algo así como: clic... clic-clic... clic... clic... clic... clic... clic-clic... clic..., etc., igual que como han escuchado al operar a un contador Geiger. Si contamos los clics que ocurren en un tiempo suficientemente largo —digamos por varios minutos— y volvemos a contar durante otro periodo igual, encontramos que los dos números son muy semejantes. Así podemos hablar de la *rapidez promedio* en la que se escuchan los clics (tantos clics por minuto en promedio).

Cuando movemos el detector, la *rapidez* con que aparecen los clics es mayor o menor, pero el tamaño (el volumen) de cada clic es siempre el mismo. Si bajamos la temperatura del alambre en el cañón, la rapidez de los clics disminuye, pero cada clic suena igual. Notamos también que si colocamos dos detectores separados en la barrera, se oye un clic en uno o en el otro, pero nunca en los dos al mismo tiempo (excepto que, de vez en cuando, si hay dos clics muy próximos en tiempo, nuestro oído no percibe la separación). Concluimos, por lo tanto, que lo que sea que llega a la barrera lo hace en “paquetes”. Todos los “paquetes” son del mismo tamaño: sólo llegan “paquetes” completos, y llegan de uno a la vez a la barrera. Diremos que: “Los electrones siempre llegan en paquetes idénticos”.

Del mismo modo que en nuestro experimento con balas, ahora podemos dar experimentalmente la respuesta a la siguiente pregunta: “¿Cuál es la probabilidad relativa de que un ‘paquete’ de electrones llegue a la barrera a diferentes distancias x del centro?” De la misma forma que antes, obtenemos la probabilidad relativa observando la rapidez de los clics, manteniendo constante el funcionamiento del cañón. La probabilidad de que los paquetes lleguen a una x específica es proporcional a la rapidez promedio de clics en x .

El resultado de nuestro experimento es la interesante curva marcada como P_{12} en la parte (c) de la figura 1-3. ¡Sí! Ésta es la manera en que se comportan los electrones.

1-5 Interferencia de ondas-electrones

Ahora tratemos de analizar la curva de la figura 1-3 para ver si podemos entender el comportamiento de los electrones. Lo primero que diríamos es que ya que llegan en paquetes, cada paquete, que podemos considerar como un electrón, ha llegado ya sea a través del agujero 1 o a través del agujero 2. Escribamos esto en forma de “Proposición”:

Proposición A: Cada electrón atraviesa *ya sea* el agujero 1 *o* el agujero 2.

Suponiendo que la Proposición A sea cierta, todos los electrones que llegan a la barrera se pueden dividir en dos clases: (1) aquellos que llegan a través del agujero 1, y (2) aquellos que llegan a través del agujero 2. De este modo, la curva que observamos debe ser la suma de los efectos de los electrones que llegan a través del agujero 1 y de los electrones que llegan a través del agujero 2. Comprobemos esta idea experimentalmente. Primero, mediremos aquellos electrones que llegan a través del agujero 1. Cerramos el agujero 2 y contamos los clics en el detector. De la rapidez de los clics, obtenemos P_1 . El resultado de la medición se muestra en la curva marcada como P_1 en la parte (b) de la figura 1-3. El resultado parece bastante razonable. De manera semejante, medimos P_2 , la distribución de probabilidad para los electrones que pasan a través del agujero 2. El resultado de esta medición también se muestra en la figura.

El resultado P_{12} obtenido con *ambos* agujeros abiertos claramente no es la suma de P_1 y P_2 , las probabilidades para cada uno de los agujeros por separado. Por analogía con nuestro experimento con ondas de agua, decimos que: “Hay interferencia”.

$$\text{Para electrones: } P_{12} \neq P_1 + P_2. \quad (1.5)$$

¿Cómo puede ocurrir tal interferencia? Quizá podríamos decir: “Bueno, esto significa, tal vez, que *no es verdad* que los paquetes pasen a través del agujero 1 o del 2, porque de ser así, las probabilidades se deberían sumar. Quizá pasen a través de los agujeros de manera más complicada. Se parten por la mitad y...” ¡Pero no! No pueden, siempre llegan en paquetes... “Bueno, quizá algunos de ellos atraviesan el 1, y después dan la vuelta a través del 2, y después dan la vuelta unas cuantas veces más, o por otro camino más complicado... entonces al cerrar el agujero 2, cambiamos la posibilidad de que un electrón que *parte* del agujero 1 llegue finalmente a la barrera...” ¡Pero noten! Hay algunos puntos a los que llegan muy pocos electrones cuando *ambos* agujeros están abiertos, pero que reciben muchos electrones si cerramos un agujero; así *cerrar* un agujero *aumentó* la cantidad del otro. Noten, sin embargo, que en el centro del patrón, P_{12} es más de dos veces mayor que $P_1 + P_2$. Es como si cerrar un agujero *disminuyera* el número de electrones que pasan a través del otro agujero. Parece difícil explicar *ambos* efectos proponiendo que los electrones recorren caminos complicados.

Esto es bastante misterioso. Y cuanto más piensen en ello parecerá más misterioso. Muchas ideas se han desarrollado para tratar de explicar la curva P_{12} en términos de electrones individuales que recorren caminos complicados a través de los agujeros. Ninguna de ellas ha tenido éxito. Ninguna de ellas puede dar la curva apropiada para P_{12} en función de P_1 y P_2 .

Aun así, lo que es bastante sorprendente, las *matemáticas* para relacionar P_1 y P_2 con P_{12} son extremadamente simples. Pues P_{12} es muy parecida a la curva I_{12} de la figura 1-2, y *ésa* fue simple. Lo que ocurre en la barrera se puede describir mediante dos números complejos, que llamamos ϕ_1 y ϕ_2 (éstos son funciones de x , naturalmente). El cuadrado de la norma de ϕ_1 produce el efecto cuando sólo el agujero 1 está abierto. Es decir, $P_1 = |\phi_1|^2$. El efecto con sólo el agujero 2 es producido por ϕ_2 de la misma manera. Es decir, $P_2 = |\phi_2|^2$. Y el efecto combinado de los dos agujeros es entonces $P_{12} = |\phi_1 + \phi_2|^2$. ¡Las *matemáticas* son las mismas que teníamos para las ondas de agua! (Es difícil ver cómo se puede obtener un resultado tan simple, a partir de un juego complicado de electrones moviéndose hacia adelante y hacia atrás a través de la placa en trayectorias extrañas.)

Concluimos lo siguiente: los electrones llegan en paquetes, como partículas, y la probabilidad de llegada de estos paquetes se ordena como la distribución para la intensidad de una onda. Es en este sentido que un electrón se comporta “algunas veces como una partícula y otras como una onda”.

Por cierto, cuando estábamos tratando con ondas clásicas definimos la intensidad como el valor medio en tiempo del cuadrado de la amplitud de onda, y usamos números complejos como un truco matemático para simplificar el análisis. Pero en mecánica cuántica resulta que las amplitudes *deben* ser representadas mediante números complejos. Las partes reales solas no bastarían. Ésta es una cuestión técnica, por el momento, ya que las fórmulas son muy semejantes.

Debido a que la probabilidad de llegada a través de ambos agujeros es tan sencilla, aunque no sea igual a $(P_1 + P_2)$, no hay nada más que añadir. Pero hay una gran cantidad de sutilezas envueltas en el hecho de que la naturaleza sí funciona de esta forma. Nos gustaría ilustrar algunas de estas sutilezas para ustedes en este momento. En primer lugar, puesto que el número que llega a un punto específico *no* es igual al número que llega a través de 1 más el número que llega a través de 2, como habíamos concluido de la Proposición A, sin duda alguna debemos concluir que *la Proposición A es falsa*. No es verdad que los electrones pasen *ya sea* a través del agujero 1 o del 2. Pero podemos comprobar esta conclusión con otro experimento.

1-6 Espiando a los electrones

Ahora probaremos el siguiente experimento. A nuestro aparato de electrones le añadimos una fuente de luz muy fuerte, colocada detrás de la pared y entre los dos agujeros, como se muestra en la figura 1-4. Sabemos que las cargas eléctricas dispersan la luz. Así que cuando un electrón pase, como quiera que lo haga, en su camino hacia el detector, éste dispersará algo de luz hacia nuestros ojos, y podremos *ver* por dónde va el electrón. Si, por ejemplo, un electrón siguiera el camino del agujero 2 que se muestra en la figura 1-4, deberíamos ver un destello de luz procedente de alrededor del lugar marcado como A en la figura. Si un electrón pasa a través del agujero 1, esperaríamos ver un destello alrededor del agujero superior. Si sucediera que vemos luz en los dos lugares al mismo tiempo, debido a que el electrón se divide en dos... ¡Hagamos el experimento!

Esto es lo que observamos: *cada vez* que oímos un clic en nuestro detector de electrones (en la barrera), *también vemos* un destello de luz *ya sea* cerca del agujero 1 *o* cerca del 2, pero *¡nunca* en ambos al mismo tiempo! Y obtenemos el mismo resultado sin importar dónde coloquemos al detector. De esta observación concluimos que cuando miramos a los electrones, encontramos que éstos pasan *ya sea* a través de un agujero *o* a través del otro. De manera experimental, la Proposición A es necesariamente verdadera.

¿Cuál es, entonces, el error en nuestro argumento en contra de la Proposición A? ¿Por qué *no es* P_{12} exactamente igual a $P_1 + P_2$? ¡Volvamos al experimento! Sigámosle la pista a los electrones y averigüemos lo que están haciendo. Para cada posición (un lugar x) del detector contaremos a los electrones que lleguen y *también* nos enteraremos por cuál agujero pasaron, observando los destellos.

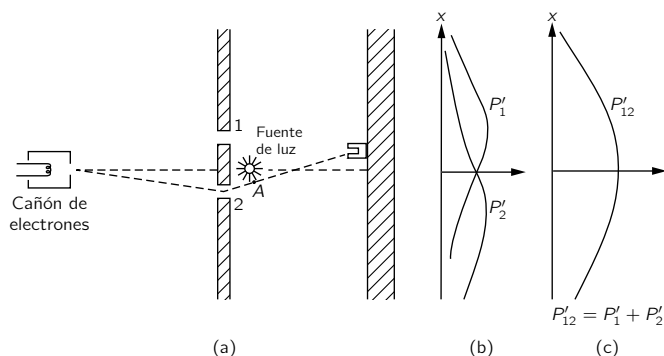


Figura 1-4. Un experimento diferente con electrones.

Podemos registrar lo que ocurre de la siguiente forma: cuando escuchemos un “clic”, lo anotaremos en la columna 1 si vemos el destello cerca del agujero 1, y si vemos el destello cerca del agujero 2, lo anotaremos en la columna 2. Cada electrón que llegue es anotado en una de estas dos clases: aquellos que lleguen a través de 1 y aquellos que llegan a través de 2. Del número registrado en la columna 1 obtenemos la probabilidad P'_1 de que un electrón llegue al detector a través del agujero 1; y del número registrado en la columna 2 obtenemos P'_2 , la probabilidad de que un electrón llegue al detector a través del agujero 2. Si repetimos estas mediciones para muchos valores de x , obtenemos las curvas para P'_1 y P'_2 mostradas en la parte (b) de la figura 1-4.

Bueno, ¡esto no es muy sorprendente! Obtenemos para P'_1 algo muy similar a lo que obtuvimos antes para P_1 cuando tapamos el agujero 2; y P'_2 es similar a lo que obtuvimos al tapar el agujero 1. Así que *no* hay nada tan complicado como pasar a través de ambos agujeros. Cuando los observamos, los electrones pasan exactamente como esperaríamos que lo hicieran. Ya sea que los agujeros estén cerrados o abiertos, aquellos que vemos pasar por el agujero 1 se distribuyen de la misma forma sin importar que el agujero 2 esté abierto o cerrado.

¡Esperemos un momento! ¿Qué obtenemos *ahora* para la probabilidad *total*, la probabilidad de que un electrón llegue al detector por cualquier camino? Ya tenemos esa información. Sólo aparentamos que no hemos visto los destellos de luz, y agrupamos los clics del detector que habíamos separado en dos columnas. Sólo *debemos sumar* los números. Para la probabilidad de que un electrón llegue a la barrera pasando por *cualquier* agujero, encontramos que $P'_{12} = P'_1 + P'_2$. Por lo tanto, aunque hemos logrado ver por cuál agujero pasan nuestros electrones, ya no obtenemos la antigua curva de interferencia P_{12} , sino una nueva, P'_{12} ; ¡que no presenta interferencia! Si apagamos la luz se vuelve a obtener P_{12} .

Debemos concluir que *cundo observamos a los electrones*, la distribución de los mismos en la pantalla es diferente que cuando no los observamos. ¿Tal vez encender nuestra luz es lo que altera las cosas? Debe ser que los electrones son muy delicados, y la luz, cuando se dispersa en los electrones, les da una sacudida que cambia su movimiento. Sabemos que el campo eléctrico de la luz que actúa sobre una carga ejercerá una fuerza sobre ella. Así, tal vez *debemos* esperar que el movimiento cambie. De todas formas, la luz ejerce una gran influencia sobre los electrones. Al intentar “espiar” a los electrones hemos cambiado su movimiento. O sea, la sacudida dada al electrón cuando el fotón es dispersado por éste, es tal que hace cambiar el movimiento del electrón lo suficiente como para que *pudiera* haber ido hacia donde P_{12} es máxima en lugar de llegar a donde P_{12} es mínima; es por esto que ya no vemos los efectos ondulantes de la interferencia.

Pueden estar pensando: “¡No usen una fuente tan brillante! ¡Disminuyan el brillo! Las ondas de luz serán entonces más débiles y no alterarán tanto a los electrones. Ciertamente, disminuyendo más y más la luz, eventualmente la onda será bastante débil que tendrá un efecto despreciable”. Muy bien. Probémoslo. Lo primero que observamos es que los destellos de luz dispersados por los electrones cuando pasan *no* se hacen más débiles. *El destello es siempre del mismo tamaño*. Lo único que sucede cuando la luz disminuye es que algunas veces oímos un “clic” en el detector pero *no vemos ningún destello*. El electrón ha pasado sin ser “visto”. Lo que observamos es que la luz *también* se comporta como los electrones; *sabíamos* que era “ondulante”, pero ahora nos damos cuenta de que también es como un “paquete”. Siempre llega —o se dispersa— en paquetes que llamamos “fotones”. Cuando disminuimos la *intensidad* de la fuente de luz no cambiamos el *tamaño* de los fotones, sino la *rapidez* con que son emitidos. *Esto* explica por qué, cuando la fuente es poco intensa, algunos electrones pasan sin ser vistos. No había ningún fotón cerca cuando pasó el electrón.

Todo esto es un poco desalentador. Si es verdad que siempre que “vemos” al electrón, vemos un destello del mismo tamaño, entonces los electrones que vemos son *siempre* los que fueron alterados. Probemos el experimento con una luz muy tenue de cualquier manera. Ahora cuando oigamos un clic en el detector lo marcaremos en tres columnas: en la columna (1) aquellos electrones vistos en el agujero 1, en la columna (2) los electrones vistos en el agujero 2, y en la

columna (3) aquellos electrones que no se vean. Cuando colectamos nuestros datos (calculando las probabilidades), encontramos estos resultados: aquellos “vistos en el agujero 1” tienen una distribución como P'_1 ; aquellos “vistos en el agujero 2” tienen una distribución como P'_2 (por lo que “aquellos vistos ya sea en el agujero 1 o en el 2” tienen una distribución como P'_{12}); y aquellos “no vistos en absoluto” tienen una distribución “ondulante” ¡justo como P_{12} en la figura 1-3! ¡Si los electrones no se ven, tenemos interferencia!

Esto es comprensible. Cuando no vemos al electrón, ningún fotón lo altera, y cuando sí lo vemos, un fotón lo ha alterado. Existe siempre la misma alteración ya que todos los fotones de la luz producen efectos del mismo tamaño, y el efecto de los fotones al ser dispersados es suficiente para anular cualquier efecto de interferencia.

¿Hay algún modo en que podamos ver a los electrones sin alterarlos? Aprendimos en un capítulo anterior que el momentum que lleva un “fotón” es inversamente proporcional a su longitud de onda ($p = h/\lambda$). Seguramente la sacudida dada al electrón cuando el fotón se dispersa hacia nuestro ojo depende del momentum que lleva el fotón. ¡Ajá! Si quiéramos alterar a los electrones sólo un poco no deberíamos haber disminuido la *intensidad* de la luz, deberíamos haber disminuido su *frecuencia* (que es lo mismo que aumentar su longitud de onda). Usemos luz de un color más rojo. Incluso podríamos usar luz infrarroja, u ondas de radio (como un radar), y “ver” por dónde va el electrón con la ayuda de algún equipo que pueda “ver” la luz de estas longitudes de onda más grandes. Si utilizamos luz “más suave” quizá evitemos alterar tanto a los electrones.

Intentemos el experimento con ondas más largas. Continuaremos repitiendo nuestro experimento cada vez con luz de una longitud de onda más grande. Al principio parece que nada cambia. Los resultados son los mismos. Luego sucede algo terrible. Recuerden que cuando discutimos el microscopio hicimos notar que, debido a la *naturaleza ondulatoria* de la luz, hay una limitación en cuán cerca pueden estar dos manchas y aún ser vistas como dos manchas separadas. Esta distancia es del orden de la longitud de onda de la luz. Ahora, cuando hacemos la longitud de onda más grande que la distancia entre nuestros agujeros, vemos un *gran* destello borroso cuando la luz es dispersada por los electrones. ¡Ya no podemos decir por cuál agujero pasó el electrón! ¡Sólo sabemos que pasó por algún lugar! Y es justo con luz de este color que notamos que las sacudidas dadas al electrón son lo suficientemente pequeñas para que P'_{12} comience a parecerse a P_{12} —que empecemos a obtener algún efecto de interferencia—. Y es sólo para longitudes de onda mucho más grandes que la separación de los dos agujeros (cuando no tenemos la más mínima oportunidad de decir por dónde pasó el electrón) como la alteración ocasionada por la luz se hace lo suficientemente pequeña para obtener de nuevo la curva P_{12} que se muestra en la figura 1-3.

En nuestro experimento encontramos que es imposible acomodar la luz de manera que podamos decir por cuál agujero atravesó el electrón, y al mismo tiempo no se altere el patrón. Heisenberg sugirió que las nuevas leyes de la naturaleza sólo podrían ser consistentes si existiesen algunas limitaciones básicas en nuestras capacidades experimentales que no hubieran sido reconocidas antes. Él propuso, como principio general, su *principio de incertidumbre*, que podemos enunciar en términos de nuestro experimento de la siguiente manera: “Es imposible diseñar un aparato para determinar por cuál agujero pasa el electrón, y que al mismo tiempo no altere a los electrones lo suficiente como para destruir el patrón de interferencia”. Si un aparato es capaz de determinar por cuál agujero pasa el electrón, *no puede* ser tan delicado como para que no altere el patrón de un modo fundamental. Nadie ha encontrado jamás (o aun imaginado) una manera de evitar el principio de incertidumbre. Por lo que debemos suponer que describe una característica básica de la naturaleza.

La teoría completa de la mecánica cuántica que ahora usamos para describir a los átomos y, en realidad, a toda la materia, depende de la validez del principio de incertidumbre. Puesto que la mecánica cuántica es una teoría muy exitosa, nuestra creencia en el principio de incertidumbre es reforzada. Pero si alguna vez se encontrase un medio para “vencer” al principio de incertidumbre, la mecánica

cuántica daría resultados inconsistentes y tendría que ser descartada como teoría válida de la naturaleza.

“Bueno”, dirán ustedes, “¿qué pasa con la Proposición A? ¿Es verdad, o *no* es verdad, que el electrón pasa ya sea a través del agujero 1 o a través del agujero 2?” La única respuesta que se puede dar es que hemos encontrado, por medio de experimentos, que hay una cierta manera especial en la que debemos pensar para evitar caer en inconsistencias. Lo que debemos decir (para evitar hacer predicciones erróneas) es lo siguiente. Si alguien observa los agujeros o, de manera más precisa, si tiene un aparato capaz de determinar si los electrones atraviesan el agujero 1 o el 2, entonces *puede* decir que atraviesan el agujero 1 o el 2. *Pero*, cuando *nadie* intenta decir por qué camino pasa el electrón, cuando no hay nada en el experimento que altere a los electrones, entonces *nadie* puede decir que un electrón pasó por el agujero 1 o por el 2. Si alguien lo dice, y comienza a hacer deducciones a partir de su afirmación, cometerá errores en el análisis. Ésta es la cuerda floja lógica sobre la que debemos caminar si deseamos describir a la naturaleza con éxito.

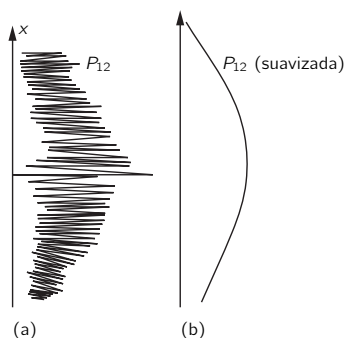


Figura 1-5. Patrón de interferencia con balas: (a) real (esquemático), (b) observado.

Si el movimiento de toda la materia —así como el de los electrones— se debe describir en términos de ondas, ¿qué pasa con las balas de nuestro primer experimento? ¿Por qué no vimos un patrón de interferencia ahí? Sucede que para las balas las longitudes de onda eran tan pequeñas, que los patrones de interferencia se volvieron muy tenues. Tan tenues, de hecho, que ningún detector de tamaño finito podría distinguir los diversos máximos y mínimos. Lo que hemos visto fue sólo una especie de promedio, que es la curva clásica. En la figura 1-5 hemos tratado de indicar esquemáticamente lo que sucede con los objetos a gran escala. La parte (a) de la figura muestra la distribución de probabilidad que se podría predecir para las balas, usando la mecánica cuántica. Las oscilaciones rápidas se supone que representan al patrón de interferencia que se obtiene para ondas con longitud de onda muy pequeña. Cualquier detector físico, sin embargo, atraviesa varias oscilaciones de la curva de probabilidad, y entonces las mediciones producen la curva suave mostrada en la parte (b) de la figura.

1-7 Los primeros principios de la mecánica cuántica

Ahora escribiremos un resumen de las principales conclusiones obtenidas de nuestros experimentos. Sin embargo, plantearemos los resultados en una forma que los haga válidos para una clase general de tales experimentos. Podemos escribir nuestro resumen de una manera más simple si primero definimos un “experimento ideal” como aquel en que no existen influencias extrañas inciertas, es decir, no hay vibraciones u otras cosas que no podamos tomar en cuenta. Seríamos más precisos si dijésemos: “Un experimento ideal es aquel en que todas las condiciones iniciales y finales del experimento están completamente especificadas”. Lo que llamaremos “un evento” es, en general, solamente un conjunto específico de condiciones iniciales y finales. (Por ejemplo: “un electrón sale del cañón, llega al detector y no sucede nada más”.) Ahora nuestro resumen.

Resumen

- (1) La probabilidad de un evento en un experimento ideal está dada por el cuadrado de la norma de un número complejo ϕ que se llama la amplitud de probabilidad:

$$\begin{aligned} P &= \text{probabilidad,} \\ \phi &= \text{amplitud de probabilidad,} \\ P &= |\phi|^2. \end{aligned} \tag{1.6}$$

- (2) Cuando un evento puede ocurrir de varias maneras, la amplitud de probabilidad para el evento es la suma de las amplitudes de probabilidad para cada alternativa considerada de manera independiente. Hay interferencia:

$$\begin{aligned} \phi &= \phi_1 + \phi_2, \\ P &= |\phi_1 + \phi_2|^2. \end{aligned} \tag{1.7}$$

- (3) Si se hace un experimento en el cual se puede determinar qué alternativa ocurre en realidad, la probabilidad del evento es la suma de las probabilidades de cada alternativa. Se pierde la interferencia:

$$P = P_1 + P_2. \quad (1.8)$$

Aún podríamos preguntar: “¿Cómo funciona esto? ¿Cuál es la maquinaria que hay detrás de la ley?” Nadie ha encontrado maquinaria alguna detrás de la ley. Nadie puede “explicar” más allá de lo que nosotros hemos “explicado”. Nadie les dará una representación más detallada de la situación. No tenemos ninguna idea acerca de algún mecanismo más básico a partir del cual se puedan deducir estos resultados.

Nos gustaría hacer notar una diferencia muy importante entre la mecánica clásica y la mecánica cuántica. Hemos estado hablando de la probabilidad de que un electrón llegue en una circunstancia determinada. Hemos dicho que en nuestro arreglo experimental (o aun en el mejor posible) sería imposible predecir exactamente lo que va a suceder. ¡Sólo podemos predecir las posibilidades! Esto significaría, si fuese verdad, que la física ha renunciado a tratar de predecir de manera exacta lo que sucederá en una circunstancia dada. ¡Sí!, la física *ha* renunciado. *No sabemos cómo predecir lo que sucedería en una circunstancia dada*, y ahora creemos que es imposible —que lo único que se puede predecir es la probabilidad para diferentes eventos—. Se debe reconocer que esto es una reducción de nuestro ideal anterior de entender la naturaleza. Puede ser un paso atrás, pero nadie ha encontrado una manera para evitarlo.

Ahora haremos algunas observaciones sobre una sugerencia que se ha hecho algunas veces tratando de evitar la descripción que hemos dado: “Quizá el electrón tenga alguna clase de mecanismos internos —algunas variables internas— acerca de los cuales no sabemos nada aún. Quizá sea por eso que no podemos predecir lo que sucederá. Si pudiésemos observar más de cerca al electrón, podríamos decir a dónde terminará”. Hasta donde sabemos, eso es imposible. Aun así tendríamos dificultades. Supongamos que dentro del electrón hay ciertos mecanismos que determinan dónde terminará. Estos mecanismos *también* deben determinar por cuál agujero va a pasar. Pero no debemos olvidar que lo que está dentro del electrón no debe depender de lo que *nosotros* hagamos, y en particular de que nosotros abramos o cerremos uno de los agujeros. Así que si un electrón, antes de empezar, ya ha decidido (a) cuál agujero va a utilizar, y (b) dónde va a terminar, deberíamos encontrar P_1 para los electrones que han elegido el agujero 1, P_2 para los que han elegido el agujero 2, y *necesariamente* la suma $P_1 + P_2$ para los que llegan a través de los dos agujeros. Parece ser que no existe otra alternativa. Pero hemos comprobado experimentalmente que éste no es el caso. Y nadie ha encontrado la manera de resolver este acertijo. Por lo tanto, actualmente debemos limitarnos a calcular probabilidades. Decimos “actualmente”, aunque sospechamos, con muchísimo fundamento, que así será para siempre —que es imposible resolver este acertijo— y que ésta es la manera en que la naturaleza realmente *es*.

1-8 El principio de incertidumbre

Heisenberg estableció de esta manera el principio de incertidumbre originalmente: si hacen una medición sobre cualquier objeto, y pueden determinar la componente x de su momentum con una incertidumbre Δp , no pueden, al mismo tiempo, conocer su posición x con una exactitud mayor que $\Delta x \geq \hbar/2\Delta p$, donde $\hbar$ es un número fijo determinado por la naturaleza. Se llama la “constante de Planck reducida”, y vale aproximadamente 1.05×10^{-34} joule-segundos. Las incertidumbres en la posición y el momentum de una partícula en algún instante deben tener un producto mayor que la mitad de la constante de Planck reducida. Éste es un caso especial del principio de incertidumbre que se enunció antes con mayor generalidad. El enunciado más general era que no se puede diseñar un equipo de modo alguno para determinar cuál de dos alternativas se lleva a cabo, sin, al mismo tiempo, destruir el patrón de interferencia.

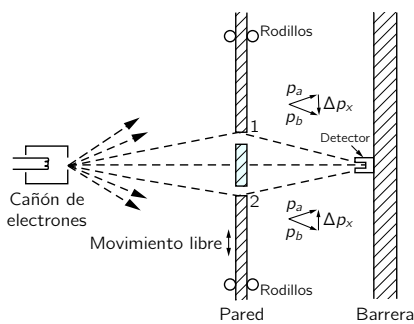


Figura 1-6. Un experimento en el cual se mide el retroceso de la pared.

Mostremos, para un caso particular, que el tipo de relación dada por Heisenberg debe ser verdadera a fin de evitar dificultades. Imaginemos una modificación del experimento de la figura 1-3, en la cual la pared con los agujeros es una placa montada sobre rodillos, de tal forma que puede moverse hacia arriba o hacia abajo libremente (en la dirección x), como se muestra en la figura 1-6. Al observar el movimiento de la placa con cuidado, podemos tratar de decir por cuál agujero pasa un electrón. Imaginemos lo que ocurre cuando el detector está colocado en $x = 0$. Esperaríamos que un electrón que pasa por el agujero 1 fuera desviado hacia abajo por la placa para alcanzar el detector. Como la componente vertical del momentum del electrón ha cambiado, la placa debe retroceder con un momentum igual en la dirección opuesta. La placa sufrirá un empujón hacia arriba. Si el electrón pasara por el agujero inferior, la placa sentiría un empujón hacia abajo. Es claro que, para cada posición del detector, el momentum recibido por la placa tendrá un valor diferente para cuando se atraviesa el agujero 1 que cuando se atraviesa el agujero 2. ¡Eso es! Sin alterar al electrón *en lo más mínimo*, sólo observando la *placa*, podemos decir qué camino utiliza el electrón.

Pero para poder hacer esto, es necesario conocer cuál es el momentum de la placa antes de que el electrón la atraviese. Así cuando medimos el momentum después de que pasó el electrón, podemos calcular cuánto ha cambiado el momentum de la placa. Pero recuerden, según el principio de incertidumbre no podemos saber al mismo tiempo la posición de la placa con una exactitud arbitraria. Pero si no sabemos exactamente *dónde* se encuentra la placa, no podemos decir de manera precisa dónde se encuentran los dos agujeros. Estarán en un lugar diferente para cada electrón que pase. Esto significa que el centro de nuestro patrón de interferencia tendrá una ubicación diferente para cada electrón. Las ondulaciones del patrón de interferencia desaparecerán. Mostraremos cuantitativamente en el próximo capítulo que si determinamos el momentum de la placa con suficiente exactitud para determinar, por las mediciones del retroceso, qué agujero se usó, entonces la incertidumbre en la posición x de la placa, de acuerdo con el principio de incertidumbre, será suficiente para desviar el patrón observado en el detector hacia arriba y hacia abajo en la dirección x , tanto como la distancia de un máximo a su mínimo más próximo. Tal desviación aleatoria es suficiente como para desvanecer el patrón por lo que no se observa ninguna interferencia.

El principio de incertidumbre “protege” a la mecánica cuántica. Heisenberg comprendió que, si fuese posible medir el momentum y la posición de manera simultánea con una mayor precisión, la mecánica cuántica colapsaría. Por lo que propuso que eso debe ser imposible. Luego, muchos se sentaron y trataron de encontrar una manera de hacerlo, y nadie pudo encontrar la manera de medir la posición y el momentum de cualquier objeto —una pantalla, un electrón, una bola de billar, cualquier cosa— con mayor exactitud. La mecánica cuántica mantiene su peligrosa, pero aún correcta, existencia.

Relación entre los puntos de vista de ondas y de partículas

2-1 Amplitudes de onda de probabilidad

En este capítulo estudiaremos la relación entre los puntos de vista de ondas y de partículas. Ya sabemos, por lo visto en el último capítulo, que ni el punto de vista de ondas ni el de partículas son correctos. Siempre nos gustaría presentar las cosas en forma exacta, o por lo menos con suficiente precisión como para que no sea necesario hacer cambios cuando aprendamos más —puede que se aumente, ¡pero no que cambie!—. Pero cuando hablamos de la imagen de onda o de la imagen de partícula, ambas son aproximadas y ambas cambiarán. Por lo tanto, lo que aprendamos en este capítulo no será preciso en cierto sentido; trataremos con algunos argumentos medio intuitivos, que se harán más precisos después. Ciertas cosas cambiarán un poco cuando las interpretemos correctamente en la mecánica cuántica. Lo hacemos de este modo para que ustedes puedan tener una impresión cualitativa de algunos fenómenos cuánticos antes de que entremos en los detalles matemáticos de la mecánica cuántica. Además, todas nuestras experiencias son con ondas y con partículas, por lo cual resulta muy práctico usar las ideas de onda y de partícula para obtener una cierta comprensión de lo que ocurre en circunstancias determinadas, antes de que veamos las matemáticas completas de las amplitudes cuánticas. Trataremos de indicar los puntos más débiles a medida que avancemos, pero una gran parte de cuanto veremos será bastante correcto —es simplemente cuestión de interpretación—.

Antes que nada, sabemos que la nueva manera de representar al mundo en la mecánica cuántica —el nuevo marco de referencia— es dar una amplitud para cada evento que puede ocurrir, y si el evento involucra la recepción de una partícula, podemos dar la amplitud para encontrar a esa partícula en lugares diferentes y en momentos distintos. La probabilidad de encontrar a la partícula es proporcional al cuadrado de la norma de la amplitud. En general, la amplitud para encontrar una partícula en lugares y momentos diferentes varía con la posición y el tiempo.

En algunos casos especiales, puede ser que la amplitud varíe de manera sinusoidal en el espacio y en el tiempo, como $e^{i(\omega t - \mathbf{k} \cdot \mathbf{r})}$, donde $\mathbf{r}$ es el vector de posición desde cierto origen. (No olviden que estas amplitudes son números complejos, no números reales.) Esta amplitud varía con una frecuencia ω y un número de onda $\mathbf{k}$ definidos. Entonces, resulta que esto corresponde a una situación límite clásica, en la que pensamos que tenemos una partícula cuya energía E es conocida, y está relacionada con la frecuencia mediante

$$E = \hbar\omega, \quad (2.1)$$

y cuyo momentum $\mathbf{p}$ también es conocido, y está relacionado con el número de onda mediante

$$\mathbf{p} = \hbar\mathbf{k}. \quad (2.2)$$

(El símbolo $\hbar$ representa al número h dividido entre 2π ; $\hbar = h/2\pi$.)

Esto significa que la idea de una partícula está limitada. La idea de una partícula —su ubicación, su momentum, etc.— que tanto usamos, es en cierto modo poco satisfactoria. Por ejemplo, si la amplitud para encontrar una partícula en diferentes lugares está dada por $e^{i(\omega t - \mathbf{k} \cdot \mathbf{r})}$, cuya norma al cuadrado es constante, significaría que la probabilidad para encontrar una partícula es la misma en todos

2-1 Amplitudes de onda de probabilidad

2-2 Medición de la posición y el momentum

2-3 Difracción de cristales

2-4 El tamaño de un átomo

2-5 Niveles de energía

2-6 Implicaciones filosóficas

Nota: Este capítulo es casi exactamente igual que el capítulo 38 del volumen I.

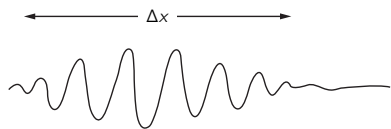


Figura 2-1. Un tren de ondas de longitud Δx .

los puntos. Esto significa que no sabemos *dónde* está —puede estar en cualquier parte—, hay una gran incertidumbre en su ubicación.

Por otra parte, si la posición de una partícula se conoce más o menos bien, y podemos predecirla con bastante precisión, entonces la probabilidad para encontrarla en diferentes lugares debe estar confinada a cierta región, cuya longitud llamamos Δx . Fuera de esta región, la probabilidad es cero. Ahora bien, esta probabilidad es el cuadrado de la norma de una amplitud, y si el cuadrado de la norma es cero, entonces la amplitud también es cero, de modo que tenemos un tren de ondas cuya longitud es Δx (figura 2-1) y la longitud de onda (la distancia entre nodos de las ondas en el tren) de ese tren de ondas es lo que corresponde al momentum de la partícula.

Aquí nos encontramos con algo extraño acerca de las ondas; algo muy simple que, estrictamente, no tiene nada que ver con la mecánica cuántica. Es algo que cualquiera que trabaja con ondas, aunque no sepa mecánica cuántica, conoce: a saber, *no podemos definir una longitud de onda única para un tren de ondas corto*. Tal tren de ondas *tiene* una longitud de onda definida; hay una indefinición en el número de onda que está relacionada con la longitud finita del tren, por lo que hay una indefinición en el momentum.

2-2 Medición de la posición y el momentum

Consideremos dos ejemplos de esta idea —para ver la razón de que haya una incertidumbre en la posición y/o en el momentum, si la mecánica cuántica es correcta—. También hemos visto antes que si no fuera así —si fuera posible medir simultáneamente la posición y el momentum de cualquier cosa— tendríamos una paradoja; somos afortunados en no tener esa paradoja, y el hecho de que esa incertidumbre surja de manera natural de la imagen de onda demuestra que todo es mutuamente consistente.

Tenemos aquí un ejemplo que muestra la relación entre la posición y el momentum en una situación que es fácil de entender. Supongamos que tenemos una sola rejilla, y que las partículas llegan de muy lejos con una energía determinada —así que todas vienen prácticamente en forma horizontal (figura 2-2)—. Nos vamos a concentrar en las componentes verticales del momentum. Todas estas partículas tienen un cierto momentum horizontal p_0 , digamos, en sentido clásico. Así, en sentido clásico, el momentum vertical p_y , antes de que la partícula atravesase el agujero, se conoce perfectamente. La partícula no se mueve ni hacia arriba ni hacia abajo, porque llegó de una fuente que está muy lejos —y por eso el momentum vertical es cero—. Pero supongamos ahora que atravesase un agujero cuyo ancho es B . Entonces, cuando sale del agujero, conocemos la posición vertical —la posición en y — con una precisión considerable —digamos $\pm B$ —†. En otras palabras, la incertidumbre en la posición, Δy , es del orden de B . Ahora también querríamos decir, dado que sabemos que el momentum es absolutamente horizontal, que Δp_y es cero; pero eso está mal. Supimos *una vez* que el momentum era horizontal, pero ya no lo sabemos más. Antes de que las partículas pasaran por el agujero, no conocíamos sus posiciones verticales. Ahora que conocemos la posición vertical, al hacer que la partícula pase por el agujero, ¡hemos perdido nuestra información sobre el momentum vertical! ¿Por qué? De acuerdo con la teoría de ondas, hay una dispersión, o difracción, de las ondas después de pasar por la rejilla, como ocurre con la luz. Por lo tanto, hay cierta probabilidad para que las partículas que salen de la rejilla no sigan exactamente en línea recta. El patrón se dispersa por efecto de la difracción, y el ángulo de dispersión, que podemos definir como el ángulo del primer mínimo, es una medida de la incertidumbre en el ángulo final.

¿Cómo se dispersa el patrón? Decir que se ha dispersado significa que hay alguna posibilidad para que la partícula se mueva hacia arriba o hacia abajo, o sea, que tiene una componente del momentum hacia arriba o hacia abajo. Decimos *posibilidad* y *partícula* porque podemos detectar este patrón de difracción con un

† Más precisamente, el error en nuestro conocimiento de y es $\pm B/2$. Pero ahora sólo estamos interesados en la idea general, por lo que no nos preocuparemos de factores de 2.

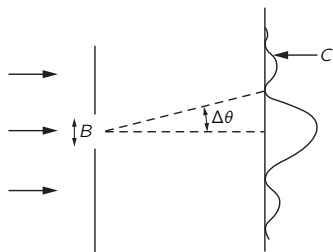


Figura 2-2. Difracción de partículas al pasar por una rejilla.

contador de partículas, y cuando el contador recibe una partícula, digamos en C en la figura 2-2, recibe la partícula *completa*, por lo que, en sentido clásico, la partícula tiene un momentum vertical, para poder ir desde la rejilla hasta C .

Para tener una idea aproximada de la dispersión del momento, el momentum vertical p_y tiene una dispersión igual a $p_0\Delta\theta$, donde p_0 es el momentum horizontal. ¿Y cuán grande es $\Delta\theta$ en el patrón ensanchado? Sabemos que el primer mínimo aparece en un ángulo $\Delta\theta$ tal que las ondas de un borde de la rejilla tienen que viajar una longitud de onda más que las provenientes del otro lado —obtuvimos este resultado antes (capítulo 30 del volumen I)—. En consecuencia, $\Delta\theta$ es λ/B , y Δp_y en este experimento es $p_0\lambda/B$. Noten que si hacemos a B más pequeño, y realizamos una medición más precisa de la posición de la partícula, el patrón de difracción se hace más ancho. Por lo tanto, cuanto más angosta hagamos la rejilla, más se ensancha el patrón, y mayor será la probabilidad para que encontremos a la partícula con un momentum lateral. Luego, la incertidumbre del momentum vertical es inversamente proporcional a la incertidumbre de y . En realidad, vemos que el producto de las dos es igual a $p_0\lambda$. Como λ es la longitud de onda y p_0 es el momentum, y de acuerdo a la mecánica cuántica, entonces la longitud de onda por el momentum es la constante de Planck h . Así que obtenemos la regla de que las incertidumbres en el momentum vertical y la posición vertical tienen un producto del orden de h :

$$\Delta y \Delta p_y \geq h/2. \quad (2.3)$$

No podemos preparar un sistema en el que sepamos la posición vertical de una partícula, y podamos predecir cómo se moverá verticalmente con mayor certeza que la dada por (2.3). Es decir, la incertidumbre en el momentum vertical debe ser mayor que $h/2\Delta y$, donde Δy es la incertidumbre en nuestro conocimiento de la posición.

Hay quienes dicen que la mecánica cuántica está por completo mal. Cuando la partícula llegó desde la izquierda, su momentum vertical era cero. Ahora que ha pasado por la rejilla, se sabe su posición. La posición y el momentum parecen conocerse con precisión arbitraria. Es completamente cierto que podemos recibir una partícula, y determinar cuál es su posición al recibirla y cuál tendría que haber sido su momentum para llegar allí. Eso es cierto, pero no es eso a lo que se refiere la relación de incertidumbre (2.3). La ecuación (2.3) se refiere a la *predictibilidad* de una situación, no a comentarios sobre el *pasado*. De nada vale decir “conocía cuál era el momentum antes de que atravesara la rejilla, y ahora conozco la posición”, porque ahora se ha perdido el conocimiento sobre el momentum. El hecho de que atravesara la rejilla, ya no nos permite predecir el momentum vertical. Estamos hablando de una teoría predictiva, y no sólo de mediciones después del hecho. Por lo que debemos hablar de lo que podemos predecir.

Ahora tratemos el asunto a la inversa. Tomemos otro ejemplo del mismo fenómeno, un poco más cuantitativamente. En el ejemplo anterior, medimos el momentum por un método clásico. Es decir, consideramos la dirección, la velocidad, los ángulos, etc., por lo que obtuvimos el momentum por un análisis clásico. Pero, como el momentum está relacionado con el número de onda, hay aún otra manera en la naturaleza para medir el momentum de una partícula —fotón o lo que sea— que no tiene un análogo clásico, porque utiliza la ecuación (2-2). Medimos las *longitudes de onda de las ondas*. Tratemos de medir el momentum de esta forma.

Supongamos que tenemos una rejilla con un gran número de líneas (figura 2-3), y enviamos un haz de partículas a la rejilla. Hemos discutido este problema a menudo: si las partículas tienen un momentum definido, entonces obtenemos un patrón muy agudo en cierta dirección, debido a la interferencia. Y también hemos hablado acerca de la precisión con que podemos determinar ese momentum, es decir, cuál es el poder de resolución de esa rejilla. En vez de deducirlo nuevamente, los referimos al capítulo 30 del volumen I, donde encontramos que la incertidumbre relativa de la longitud de onda que se puede medir con una rejilla determinada es $1/Nm$, donde N es el número de líneas de la rejilla y m es el orden del patrón de difracción. Esto es,

$$\Delta\lambda/\lambda = 1/Nm. \quad (2.4)$$

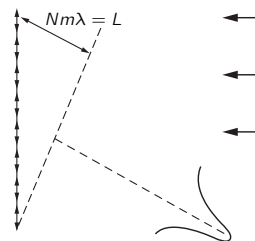


Figura 2-3. Determinación del momentum usando una rejilla de difracción.

Ahora bien, se puede reescribir la fórmula (2.4) en la forma

$$\Delta\lambda/\lambda^2 = 1/Nm\lambda = 1/L, \quad (2.5)$$

donde L es la distancia mostrada en la figura 2-3. Esta distancia es la diferencia entre la distancia total que la partícula, o la onda, o lo que sea, tiene que viajar si es reflejada en la base de la rejilla, y la distancia que tiene que viajar si es reflejada en la parte superior de la rejilla. Es decir, las ondas que forman el patrón de difracción son ondas que vienen de partes diferentes de la rejilla. Las primeras que llegan vienen de la base de la rejilla, del principio del tren de ondas, y el resto de ellas viene de partes posteriores del tren de ondas, que llegan de partes diferentes de la rejilla, hasta que finalmente llega la última, y ésta proviene de un punto del tren de ondas que está a una distancia L detrás del primer punto. Entonces, para que tengamos una línea angosta en el espectro correspondiente a un momentum definido, con una incertidumbre dada por (2.4), necesitamos un tren de ondas por lo menos de longitud L . Si el tren de ondas es demasiado corto, no estamos utilizando toda la rejilla. Las ondas que forman el espectro se están reflejando sólo en un sector muy corto de la rejilla si el tren de ondas es muy corto, por lo cual la rejilla no funcionará bien —encontraremos una dispersión angular grande—. Para obtener una dispersión más angosta, tenemos que utilizar la rejilla completa, de modo que, por lo menos en algún momento, todo el tren de ondas se disperse simultáneamente en todos los puntos de la rejilla. Por lo que el tren de ondas debe ser de longitud L para tener una incertidumbre en la longitud de onda menor que la dada por (2.5). Además,

$$\Delta\lambda/\lambda^2 = \Delta(1/\lambda) = \Delta k/2\pi. \quad (2.6)$$

Por lo tanto,

$$\Delta k = 2\pi/L, \quad (2.7)$$

donde L es la longitud del tren de ondas.

Esto significa que si tenemos un tren de ondas cuya longitud es menor que L , la incertidumbre del número de onda debe ser mayor que $2\pi/L$. O la incertidumbre de un número de onda por la longitud del tren de ondas —la llamaremos Δx por un momento— es mayor que 2π . La llamamos Δx porque es la incertidumbre en la ubicación de la partícula. Si el tren de ondas existe sólo en una longitud finita, entonces es ahí donde podríamos encontrar a la partícula, con una incertidumbre Δx . Ahora bien, esta propiedad de las ondas, que la longitud del tren de ondas por la incertidumbre en el número de onda asociado con él es por lo menos 2π , es una propiedad conocida para cualquiera que las estudie. No tiene nada que ver con la mecánica cuántica. Es simplemente que, si tenemos un tren finito, no podemos contar las ondas que hay en él en forma muy precisa.

Intentemos otra manera para entender la razón de eso. Supongamos que tenemos un tren finito de longitud L ; entonces, por la manera en que tiene de decrecer en los extremos, como en la figura 2-1, el número de ondas en la longitud L es incierto en algo así como ± 1 . Pero el número de ondas en L es $kL/2\pi$. Por lo tanto k es incierta, y obtenemos nuevamente el resultado (2.7), que es simplemente una propiedad de las ondas. Lo mismo ocurre si las ondas son en el espacio y k es el número de radianes por centímetro, y L es la longitud del tren, que si las ondas son en el tiempo y ω es el número de oscilaciones por segundo y T es la “longitud” en el tiempo que el tren de ondas tarda en pasar. Es decir, si tenemos un tren de ondas que sólo dura un tiempo finito T , entonces la incertidumbre en la frecuencia está dada por

$$\Delta\omega = 2\pi/T. \quad (2.8)$$

Hemos tratado de enfatizar que éstas son propiedades de las ondas, y que se conocen muy bien, por ejemplo, en la teoría del sonido.

La cuestión es que, en la mecánica cuántica, interpretamos al número de onda como una medida del momentum de una partícula, con la regla de que $p = \hbar k$, por lo que la relación (2.7) nos dice que $\Delta p \approx \hbar/\Delta x$. Ésta es, entonces,

una limitación de la idea clásica de momentum. (Naturalmente, ¡tiene que estar limitada de alguna manera si vamos a representar a las partículas mediante ondas!) Es agradable que hayamos encontrado una regla que nos da una idea de cuándo hay una falla en las ideas clásicas.

2-3 Difracción de cristales

A continuación, consideremos la reflexión de ondas-partículas en un cristal. Un cristal es una cosa gruesa que tiene muchos átomos similares —más tarde incluiremos algunas complicaciones— en un arreglo ordenado. La cuestión es cómo hacemos el arreglo para que obtengamos un máximo de reflexión intenso en una dirección determinada para un haz determinado de, por ejemplo, luz (rayos X), electrones, neutrones o cualquier otra cosa. Para obtener una reflexión intensa, la dispersión en todos los átomos debe estar en fase. No puede haber un número igual en fase y fuera de fase, porque las ondas se cancelarían. El modo de arreglar las cosas es encontrar las regiones de fase constante, como ya lo hemos explicado; son los planos que forman ángulos iguales con las direcciones inicial y final (figura 2-4).

Si consideramos dos planos paralelos, como en la figura 2-4, las ondas dispersadas en los dos planos estarán en fase, siempre que la diferencia en la distancia que ha viajado el frente de onda sea un múltiplo entero de longitudes de onda. Se puede ver que esta diferencia es $2d \sin \theta$, donde d es la distancia perpendicular entre los planos. Luego entonces, la condición para reflexión coherente es

$$2d \sin \theta = n\lambda \quad (n = 1, 2, \dots). \quad (2.9)$$

Si, por ejemplo, el cristal es tal que los átomos están sobre planos que obedecen la condición (2.9) con $n = 1$, habrá una reflexión intensa. Si, por otra parte, hay otros átomos de la misma naturaleza (misma densidad) a la mitad, entonces los planos intermedios también dispersarán con igual intensidad e interferirán con los otros, y no producirán ningún efecto. Por lo tanto d en (2.9) se debe referir a planos *adyacentes*; ¡no podemos tomar un plano que esté cinco capas más atrás y usar esta fórmula!

Como punto de interés, los cristales reales por lo general no son tan simples como una sola clase de átomos repetidos de cierta manera. Por el contrario, si hacemos una analogía bidimensional, son muy parecidos al papel tapiz, en el cual hay un tipo de figura que se repite por todo el papel. Por “figura” queremos decir, en el caso de átomos, algún arreglo —calcio y un carbono y tres oxígenos para el carbonato de calcio, etc.— que puede incluir un número relativamente grande de átomos. Pero cualquiera que sea, la figura se repite en un patrón. La figura básica se denomina *celda unitaria*.

El patrón básico de repetición define lo que llamamos *red cristalina*; la red cristalina se puede determinar inmediatamente al observar las reflexiones y determinar su simetría. En otras palabras, donde encontramos *alguna* reflexión, ésta determina el tipo de red, pero para determinar qué hay en cada elemento de la red, debemos considerar la *intensidad* de la dispersión en diversas direcciones. En *cuáles* direcciones hay dispersión depende del tipo de red, pero la *intensidad* de dispersión está determinada por lo que hay dentro de cada celda unitaria, y así se obtiene la estructura de los cristales.

Dos fotografías de patrones de diagramas de rayos Xs se muestran en las figuras 2-5 y 2-6; éstas muestran la dispersión en sal de roca y en mioglobina, respectivamente.

Además, un detalle interesante ocurre si el espaciamiento de los planos más cercanos es menor que $\lambda/2$. En este caso, no hay ninguna n que satisfaga (2.9). Así, si λ es mayor de dos veces la distancia entre planos adyacentes, entonces no hay patrón de difracción lateral, y la luz —o lo que sea— atravesará el material sin rebotar o perderse. Así que en el caso de la luz, donde λ es mucho mayor que el espaciamiento, pasa directamente y no hay patrón de reflexión de los planos del cristal.

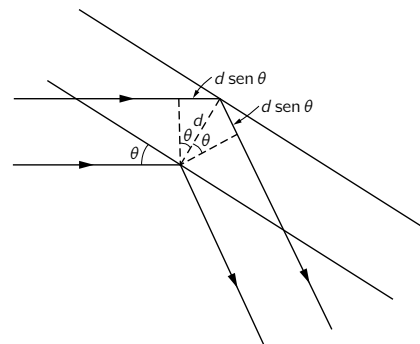


Figura 2-4. Dispersión de ondas en planos cristalinos.

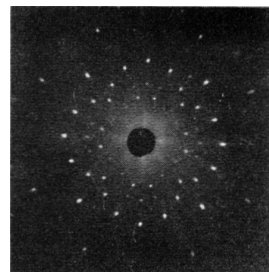


Figura 2-5

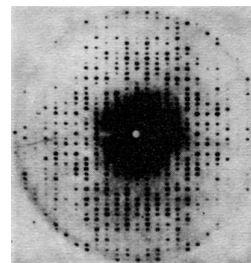


Figura 2-6

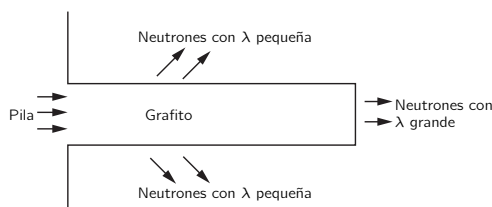


Figura 2-7. Difusión de neutrones de una pila a través de un bloque de grafito.

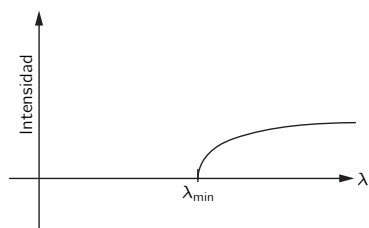


Figura 2-8. Intensidad de los neutrones que salen de la barra de grafito, en función de la longitud de onda.

Este hecho también tiene una consecuencia interesante en el caso de las pilas que hacen neutrones (éstos son evidentemente partículas, ¡pueden apostar!). Si tomamos estos neutrones y los hacemos entrar en un bloque largo de grafito, los neutrones se difunden y avanzan gradualmente (figura 2-7). Se difunden porque rebotan en los átomos, pero estrictamente, en la teoría ondulatoria, rebotan en los átomos debido a la difracción en los planos cristalinos. Ocurre que si tomamos una pieza muy larga de grafito, ¡los neutrones que salen por el extremo final tienen todos una longitud de onda larga! De hecho, si graficamos la intensidad en función de la longitud de onda, no obtenemos nada excepto para longitudes de onda más largas que un cierto mínimo (figura 2-8). En otras palabras, de este modo podemos obtener neutrones muy lentos. Sólo pasan los neutrones más lentos; éstos no son difractados ni dispersados por los planos cristalinos del grafito, prosiguen su camino como la luz a través del vidrio, y no son dispersados hacia los lados. Hay muchas otras demostraciones de la realidad de las ondas-neutrones y de otras ondas-partículas.

2-4 El tamaño de un átomo

Consideremos ahora otra aplicación de la relación de incertidumbre (2.3). No se debe tomar muy seriamente; la idea es correcta, pero el análisis no es muy preciso. La idea tiene que ver con la determinación del tamaño de los átomos, y con el hecho de que, clásicamente, los electrones radian luz y caen en espiral hasta quedarse encima del núcleo. Pero esto no puede ser cuánticamente correcto, porque entonces sabríamos dónde está cada electrón y la rapidez con que se mueve.

Supongamos que tenemos un átomo de hidrógeno, y medimos la posición del electrón; no debemos ser capaces de predecir exactamente dónde estará el electrón, porque la distribución de momentum tendría una extensión infinita. Cada vez que observemos el electrón, éste estará en alguna parte, pero tiene una amplitud para estar en lugares diferentes por lo que hay una probabilidad para encontrarlo en lugares distintos. Estos lugares no pueden estar todos en el núcleo; supondremos que hay una distribución de posiciones del orden de a . Es decir, la distancia del electrón al núcleo es por lo general alrededor de a . Determinaremos a minimizando la energía total del átomo.

La distribución de momentum es más o menos $\hbar/a$ debido a la incertidumbre, de modo que si tratamos de medir el momentum del electrón de alguna manera, como dispersando rayos X y observando el efecto Doppler de un escáner en movimiento, esperaríamos no obtengamos cero todo el tiempo —el electrón no está inmóvil— sino que el momentum debe ser del orden de $p \approx \hbar/a$. Luego, la energía cinética es aproximadamente $\frac{1}{2}mv^2 = p^2/2m = \hbar^2/2ma^2$. (En cierto sentido, esto es algo como un análisis dimensional para encontrar en qué forma la energía cinética depende de la constante de Planck reducida, de m y del tamaño del átomo. Nuestra respuesta puede variar por factores como 2, π , etc. Ni siquiera hemos definido a de manera muy precisa.) Ahora bien, la energía potencial es menos e^2 sobre la distancia al centro, es decir $-e^2/a$, donde, como se definió en el volumen I, e^2 es la carga del electrón al cuadrado dividida entre $4\pi\epsilon_0$. La cuestión es que la energía potencial se reduce si a disminuye, pero cuanto menor es a , mayor es el momentum necesario, debido al principio de incertidumbre, y por lo tanto la energía cinética es mayor. La energía total es

$$E = \hbar^2/2ma^2 - e^2/a. \quad (2.10)$$

No sabemos cuánto vale a , pero sabemos que el átomo se las va a arreglar para que la energía sea lo menor posible. Para minimizar E , derivamos respecto a a , igualamos la derivada a cero y despejamos a . La derivada de E es

$$dE/da = -\hbar^2/ma^3 + e^2/a^2, \quad (2.11)$$

y haciendo $dE/da = 0$ se obtiene para a el valor

$$\begin{aligned} a_0 &= \hbar^2/me^2 = 0.528 \text{ \AA}, \\ &= 0.528 \times 10^{-10} \text{ metros.} \end{aligned} \quad (2.12)$$

Esta distancia en particular se denomina *radio de Bohr*, y hemos aprendido que las dimensiones atómicas son del orden de *ángstroms*, lo cual es correcto. Esto es bastante bueno —en realidad es asombroso, ¡porque hasta ahora no hemos tenido ninguna base para comprender el tamaño de los átomos!—. Los átomos son completamente imposibles desde el punto de vista clásico, ya que los electrones caerían en espiral sobre el núcleo.

Si introducimos el valor (2.12) de a_0 en (2.10) resulta la energía

$$E_0 = -e^2/2a_0 = -me^4/2\hbar^2 = -13.6 \text{ eV.} \quad (2.13)$$

¿Qué significa una energía negativa? Significa que el electrón tiene menos energía cuando está en el átomo que cuando está libre. Significa que está ligado. Significa que se necesita energía para sacarlo; se necesita una energía del orden de 13.6 eV para ionizar un átomo de hidrógeno. No tenemos ninguna razón para no pensar que sea dos o tres veces mayor —o la mitad— o $(1/\pi)$ veces ese valor, porque hemos usado un argumento muy descuidado. Sin embargo, hemos hecho trampa, ¡usamos todas las constantes de modo que resulte el número correcto! Este número, 13.6 electronvolts, se denomina un *Rydberg de energía*; es la energía de ionización del hidrógeno.

Ahora entendemos por qué no atravesamos el piso. Cuando caminamos, nuestros zapatos con las masas de sus átomos empujan contra el piso con *su* masa de átomos. Para comprimir los átomos entre sí, los electrones estarían confinados en un espacio más pequeño y, por el principio de incertidumbre, sus momentums deberían ser mayores en promedio, y eso significaría una energía más alta; la resistencia a la compresión atómica es un efecto cuántico y no uno clásico. De manera clásica, esperaríamos que si acercáramos a los electrones y a los protones, la energía se reduciría aún más, y la mejor disposición de las cargas negativas y positivas en la física clásica es unas sobre otras. Esto se sabía muy bien en la física clásica y constituía un enigma debido a la existencia del átomo. Por supuesto, los científicos anteriores crearon sus propios medios para salir del aprieto —pero no importa, ¡ahora tenemos la manera *correcta* para salir del aprieto—.

Además, aunque no tenemos razón para entenderlo por el momento, en una situación donde hay muchos electrones, ocurre que tratan de mantenerse alejados unos de otros. Si un electrón ocupa un cierto espacio, otro no ocupa el mismo espacio. Para ser más precisos, hay dos lugares en el espín, de modo que dos electrones se pueden colocar uno sobre otro, uno girando en un sentido y otro en el sentido opuesto. Pero después de esto no podemos poner más ahí. Tenemos que poner a los demás en otro lugar, y ésa es la razón real de que la materia tenga resistencia. Si pudiéramos poner a todos los electrones en el mismo lugar, se condensaría aún más de lo que lo hace. El hecho de que los electrones no puedan ponerse uno sobre otro es lo que hace que las mesas y las demás cosas sean sólidas.

Es evidente que, para comprender las propiedades de la materia, tendremos que usar la mecánica cuántica y no quedarnos satisfechos con la mecánica clásica.

2-5 Niveles de energía

Hemos hablado del átomo en su condición de energía más baja posible, pero ocurre que el electrón puede hacer otras cosas. Puede saltar y oscilar con más energía, por lo que hay muchos movimientos diferentes posibles para el átomo. Según la mecánica cuántica, en una condición estacionaria, el átomo sólo puede tener energías definidas. Hagamos un diagrama (figura 2-9) representando a la energía verticalmente, y trazando una línea horizontal para cada valor permitido de la energía. Cuando el electrón está libre, es decir, cuando su energía es positiva,

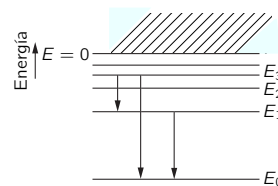


Figura 2-9. Diagrama de energía para un átomo, mostrando varias transiciones posibles.

puede tener cualquier energía; se puede mover con cualquier velocidad. Pero las energías de unión no son arbitrarias. El átomo debe tener cualquiera de un conjunto de valores permitidos, como los de la figura 2-9.

Llamemos E_0 , E_1 , E_2 , E_3 a los valores permitidos de la energía. Si un átomo está inicialmente en uno de estos “estados excitados”, E_1 , E_2 , etc., no se queda en ese estado para siempre. Tarde o temprano cae a un estado más bajo radiando energía en forma de luz. La frecuencia de la luz emitida está determinada por la conservación de la energía, más el entendimiento cuántico de que la frecuencia de la luz está relacionada con la energía de la luz por medio de la ecuación (2.1). Por lo tanto, la frecuencia de la luz liberada en una transición de la energía E_3 a la E_1 (por ejemplo) es

$$\omega_{31} = (E_3 - E_1)/\hbar. \quad (2.14)$$

Ésta es, entonces, una frecuencia característica del átomo y define una línea espectral de emisión. Otra transición posible sería de E_3 a E_0 . Ésta tendría una frecuencia diferente:

$$\omega_{30} = (E_3 - E_0)/\hbar. \quad (2.15)$$

Otra posibilidad es que, si el átomo fuera excitado al estado E_1 , éste podría caer al estado base E_0 , emitiendo un fotón de frecuencia

$$\omega_{10} = (E_1 - E_0)/\hbar. \quad (2.16)$$

La razón de que mostremos tres transiciones es que queremos señalar una relación interesante. Es fácil ver de (2.14), (2.15) y (2.16) que

$$\omega_{30} = \omega_{31} + \omega_{10}. \quad (2.17)$$

En general, si encontramos dos líneas espectrales, esperamos encontrar otra línea en la suma de las frecuencias (o en la diferencia de las frecuencias), y que todas las líneas se pueden comprender encontrando una serie de niveles tales que cada línea corresponda a la diferencia de energía de algún par de niveles. Esta coincidencia notable en las frecuencias espectrales fue observada antes de que la mecánica cuántica fuera descubierta, y se denomina el *principio de combinaciones de Ritz*. Éste es otro misterio desde el punto de vista de la mecánica clásica. No insistamos sobre el fracaso de la mecánica clásica en el dominio atómico; nos parece que ya lo hemos demostrado bastante bien.

Ya hemos hablado sobre la mecánica cuántica, y la hemos representado por amplitudes que se comportan como ondas, con ciertas frecuencias y ciertos números de onda. Veamos cómo se explica, desde el punto de vista de las amplitudes, que el átomo tenga estados de energía definidos. Esto es algo que no podemos comprender a partir de lo que se ha dicho hasta ahora, pero todos sabemos que las ondas confinadas tienen frecuencias definidas. Por ejemplo, si el sonido se confina a un tubo de órgano, o cualquier cosa parecida, entonces hay más de una manera en la que el sonido puede vibrar, pero para cada manera hay una frecuencia definida. Así, un objeto en el cual las ondas están confinadas tiene ciertas frecuencias de resonancia. En consecuencia, es una propiedad de las ondas en un espacio confinado —tema que más adelante estudiaremos en detalle con fórmulas— que éstas existan sólo para frecuencias definidas. Y como existe la relación general entre frecuencia de las amplitudes y energía, no nos sorprendamos de encontrar energías definidas asociadas con los electrones unidos a los átomos.

2-6 Implicaciones filosóficas

Consideremos brevemente algunas implicaciones filosóficas de la mecánica cuántica. Como siempre, surgen dos aspectos del problema: uno es la implicación filosófica para la física, y el otro es la extrapolación de implicaciones filosóficas en otros campos. Cuando las ideas filosóficas asociadas con la ciencia se trasladan a otros campos, generalmente se las distorsiona por completo. Por esta razón, limitaremos en lo posible nuestros comentarios a la física misma.

Antes que nada, el aspecto más interesante es la idea del principio de incertidumbre; hacer una observación afecta al fenómeno. Siempre se ha sabido que hacer observaciones afecta al fenómeno, pero la cuestión es que el efecto no se puede ignorar, o minimizar, o disminuir arbitrariamente reordenando los aparatos. Cuando buscamos un cierto fenómeno, no podemos evitar el alterarlo de alguna manera mínima, y *la alteración es necesaria para la consistencia del punto de vista*. A veces el observador era importante en la física precuántica, pero sólo de un modo trivial. Se ha propuesto el siguiente problema: si un árbol cae en un bosque y no hay nadie ahí para oírlo, ¿hace algún ruido? Claro que un árbol *real* que cae en un bosque *real* hace ruido, aunque no haya nadie ahí. Aunque nadie esté presente para oírlo, quedan otros indicios. El sonido sacudirá algunas hojas, y si fuéramos suficientemente cuidadosos, podríamos encontrar en algún lugar donde una espina rozó a una hoja y le hizo un rasguño muy ligero, que no se podría explicar a no ser que supiéramos que la hoja estaba vibrando. En cierto modo, tendríamos que admitir que se produjo un sonido. Podríamos preguntar: ¿hubo una *sensación* de sonido? No, las sensaciones tienen que ver, presumiblemente, con la conciencia. Y si las hormigas tienen conciencia, o si había hormigas en el bosque, o si el árbol tenía conciencia, no lo sabemos. Dejemos el problema de ese modo.

Otro aspecto que la gente ha enfatizado desde que se desarrolló la mecánica cuántica, es la idea de que no deberíamos hablar de las cosas que no podemos medir. (De hecho, la teoría de la relatividad también decía eso.) A no ser que algo se pueda definir por medio de una medición, no tiene cabida en una teoría. Y, como un valor preciso del momentum de una partícula localizada no se puede definir por medición, no puede, por lo tanto, formar parte de la teoría. La idea de que esto es lo que ocurría en la teoría clásica *es una posición falsa*. Es un análisis descuidado de la situación. El que no podamos *medir* el momentum y la posición en forma precisa, no significa *a priori* que *no podamos* hablar de ellos. Sólo significa que *no necesitamos* hablar de ellos. La situación en las ciencias es ésta: un concepto o una idea que no se puede medir o no se puede referir directamente al experimento, puede ser útil o no. No necesita existir en la teoría. En otras palabras, supongamos que comparamos la teoría clásica del mundo con la teoría cuántica del mundo, y supongamos que experimentalmente es verdad que podamos medir la posición y el momentum sólo en forma imprecisa. La cuestión es si las *ideas* de posición exacta de una partícula y de momentum exacto de una partícula son válidos o no. La teoría clásica admite estas ideas; la teoría cuántica no lo hace. Esto no significa por sí solo que la física clásica es incorrecta. Cuando se descubrió la nueva mecánica cuántica, las personas clásicas —que incluían a todos excepto Heisenberg, Schrödinger y Born— dijeron: “Miren, su teoría no es satisfactoria porque no puede responder ciertas preguntas como: ¿cuál es la posición exacta de una partícula?, ¿a través de qué agujero pasa? y algunas otras”. La respuesta de Heisenberg fue: “No necesito responder a esas preguntas, porque ustedes no pueden hacer ese tipo de pregunta experimentalmente”. Y es que *no tenemos* que hacerlas. Consideremos dos teorías (a) y (b); (a) contiene una idea que no se puede verificar directamente, pero que se emplea en el análisis, y la otra, (b), no contiene la idea. Si éstas no concuerdan en sus predicciones, no se podría sostener que (b) es falsa porque no puede explicar esa idea que está en (a), porque esa idea es algo que no se puede verificar directamente. Siempre es bueno saber qué ideas no se pueden verificar directamente, pero no es necesario suprimirlas. No es verdad que podamos progresar en la ciencia utilizando sólo los conceptos que están directamente sujetos a la experimentación.

En la mecánica cuántica hay una amplitud de probabilidad, un potencial y muchas construcciones conceptuales que no podemos medir directamente. El fundamento de una ciencia es su capacidad de *predecir*. Predecir significa decir qué ocurrirá en un experimento que nunca ha sido realizado. ¿Cómo podemos hacer eso? Suponiendo que sabemos lo que ocurre, independientemente del experimento. Debemos extrapolar los experimentos a una región donde no se han hecho. Debemos tomar nuestros conceptos y extenderlos a lugares donde todavía no han sido verificados. Si no hacemos esto, no tenemos predicción. Por

ello era perfectamente sensato para los físicos clásicos proseguir alegremente y suponer que la posición —que evidentemente significa algo para una pelota de béisbol— también significaba algo para un electrón. No era una estupidez. Era un procedimiento sensato. Hoy decimos que se supone que la ley de relatividad es válida para todas las energías, pero algún día puede aparecer alguien y diga lo tontos que fuimos. No sabremos dónde somos “tontos” hasta que no “salgamos de nuestra caja”. Y la única manera de saber que estamos equivocados es descubriendo *cuáles* son nuestras predicciones. Es absolutamente necesario hacer construcciones conceptuales.

Ya hemos hecho algunos comentarios sobre la incertidumbre de la mecánica cuántica. Es decir, que ahora somos incapaces de predecir lo que ocurrirá físicamente en una situación física determinada, la cual se prepara lo más cuidadosamente posible. Si tenemos un átomo que está en un estado excitado, por lo que va a emitir un fotón, no podemos decir *cuándo* emitirá el fotón. Tiene una cierta amplitud para emitir el fotón en cualquier momento, y sólo podemos predecir una probabilidad de emisión; no podemos predecir el futuro con exactitud. Esto ha dado lugar a todo tipo de tonterías y preguntas sobre el significado del libre albedrío, y a la idea de que el mundo es incierto.

Naturalmente debemos destacar que la física clásica es también indeterminada, en cierto sentido. Generalmente se piensa que esta indeterminación, el que no podamos predecir el futuro, es una cuestión cuántica de gran importancia, y se dice que explica el comportamiento de la mente, el sentimiento de libre albedrío, etc. Pero si el mundo *fuera* clásico —si las leyes de la mecánica fueran clásicas— no es del todo evidente que la mente no sentiría más o menos lo mismo. En sentido clásico, si conociéramos la posición y la velocidad de cada partícula en el mundo, o en una caja con gas, es cierto que podríamos predecir exactamente qué ocurriría. Y por lo tanto el mundo clásico es determinista. Supongamos, sin embargo, que tenemos una precisión finita y que no conocemos *exactamente* en dónde está un átomo, digamos que con un error de una mil millonésima. Entonces, a medida que este átomo sigue su camino y choca con otro átomo, y como no conociamos su posición mejor que en una mil millonésima parte, encontramos un error aún mayor en la posición después de la colisión. Y éste se amplifica, por supuesto, en la colisión siguiente, de modo que si comenzamos con sólo un error muy pequeño, éste rápidamente aumenta hasta una gran incertidumbre. Por dar un ejemplo: cuando el agua cae encima de un dique, salpica. Si estamos cerca, entonces de vez en cuando una gota caerá en nuestra nariz. Esto parece ser completamente aleatorio; no obstante, ese comportamiento se podría predecir por medio de leyes puramente clásicas. La posición exacta de las gotas depende de la ondulación precisa del agua antes de caer encima del dique. ¿Cómo? Las irregularidades más pequeñas aumentan en la caída, por lo que obtenemos un movimiento completamente aleatorio. Evidentemente, no podemos predecir la posición de las gotas a no ser que conozcamos el movimiento del agua en forma *absolutamente exacta*.

Hablando con más precisión, dada una exactitud arbitraria, por más precisa que sea, podemos encontrar un tiempo posterior como para que no podamos hacer predicciones válidas para ese tiempo. La cuestión es que ese tiempo posterior no es muy grande. No es que ese tiempo sea millones de años, si la precisión es una mil millonésima parte. En realidad, el tiempo crece de manera logarítmica con el error, y ocurre que en un tiempo bastante pequeño perdemos toda nuestra información. Si se toma una precisión de una parte en millones y millones y millones —todos los millones que queramos, mientras que nos detengamos en alguna parte—, podemos encontrar un tiempo menor que el tiempo que nos llevó determinar la precisión —después del cual ¡ya no podemos predecir qué es lo que va a ocurrir!—. Por lo tanto, no es justo decir que a partir de la libertad aparente y la indeterminación de la mente humana, tendríamos que habernos dado cuenta de que la física clásica “determinística” no habría tenido esperanza de comprenderla, y recibir a la mecánica cuántica como una liberación de un universo “completamente mecanicista”. Porque ya en la mecánica clásica había una indeterminación desde el punto de vista práctico.

Amplitudes de probabilidad

3-1 Leyes para combinar amplitudes

Cuando Schrödinger descubrió las leyes correctas de la mecánica cuántica, él escribió una ecuación que describía la amplitud para encontrar a una partícula en varios lugares. Esta ecuación era muy parecida a las ecuaciones ya conocidas por los físicos clásicos —ecuaciones que ellos habían usado para describir el movimiento del aire en una onda de sonido, la transmisión de la luz, etc.—. Por lo que, al principio de la mecánica cuántica, la mayor parte del tiempo se dedicó a resolver esta ecuación. Al mismo tiempo, se desarrolló la comprensión, en particular por Born y Dirac, de las, básicamente, nuevas ideas físicas que encierra la mecánica cuántica. Cuando la mecánica cuántica se desarrolló más, se observó que había un gran número de aspectos que no estaban directamente comprendidos en la ecuación de Schrödinger —tales como el espín del electrón y diversos fenómenos relativistas—. Tradicionalmente, todos los cursos de mecánica cuántica han comenzado del mismo modo, trazando el camino seguido en el desarrollo histórico del tema. Primero se aprende mucho sobre mecánica clásica para poder entender cómo resolver la ecuación de Schrödinger. Luego se dedica mucho tiempo a encontrar diversas soluciones. Sólo después de un detallado estudio de esta ecuación se llega a la materia más “avanzada” del espín del electrón.

Nosotros habíamos considerado, inicialmente, que la forma correcta de concluir estas lecciones de física, era mostrar cómo resolver las ecuaciones de la física clásica en situaciones complicadas —tales como la descripción de ondas acústicas en regiones cerradas, los modos de radiación electromagnética en cavidades cilíndricas, etc.—. Éste era el plan original del curso. Sin embargo, hemos decidido abandonar este plan y, en su lugar, dar una introducción a la mecánica cuántica. Hemos llegado a la conclusión de que, por costumbre, se llaman partes avanzadas de la mecánica cuántica las que son, de hecho, bastante simples. Las matemáticas necesarias son particularmente sencillas, constan de operaciones algebraicas simples sin ecuaciones diferenciales o, a lo sumo, de algunas muy sencillas. El único problema es que debemos dar el salto de ya no poder describir el comportamiento, *en detalle*, de las partículas en el espacio. Así que esto es lo que trataremos de hacer: hablaremos de lo que convencionalmente se llamarían las partes “avanzadas” de la mecánica cuántica. Pero que son, sin duda, se lo aseguramos, las partes más simples —en el sentido profundo de la palabra— así como las partes más básicas. Éste es, con toda franqueza, un experimento pedagógico; nunca se ha realizado antes, que nosotros sepamos.

En este tema tenemos, naturalmente, la dificultad de que el comportamiento cuántico de las cosas es bastante extraño. Nadie tiene una experiencia cotidiana en la que pueda apoyarse para formarse una idea vaga e intuitiva de lo que sucederá. Así que hay dos modos de presentar el tema: podríamos describir lo que puede suceder de una forma física un poco imprecisa, diciéndoles más o menos lo que sucede, sin darles las leyes precisas de todo; o podríamos, por el contrario, dar las leyes precisas en su forma abstracta. Pero entonces, a causa de las abstracciones, no sabrían de qué se tratan, en sentido físico. El último método no es adecuado, porque es completamente abstracto, y el primero deja un sentimiento incómodo, porque no sabemos exactamente lo que es verdadero y lo que es falso. No estamos seguros de cómo resolver esta dificultad. Habrán notado, de hecho, que los capítulos 1 y 2 mostraban este problema. El primer capítulo fue relativamente preciso, pero el segundo fue una descripción vaga de

3-1 Leyes para combinar amplitudes

3-2 El patrón de interferencia de dos rejillas

3-3 Dispersión en un cristal

3-4 Partículas idénticas

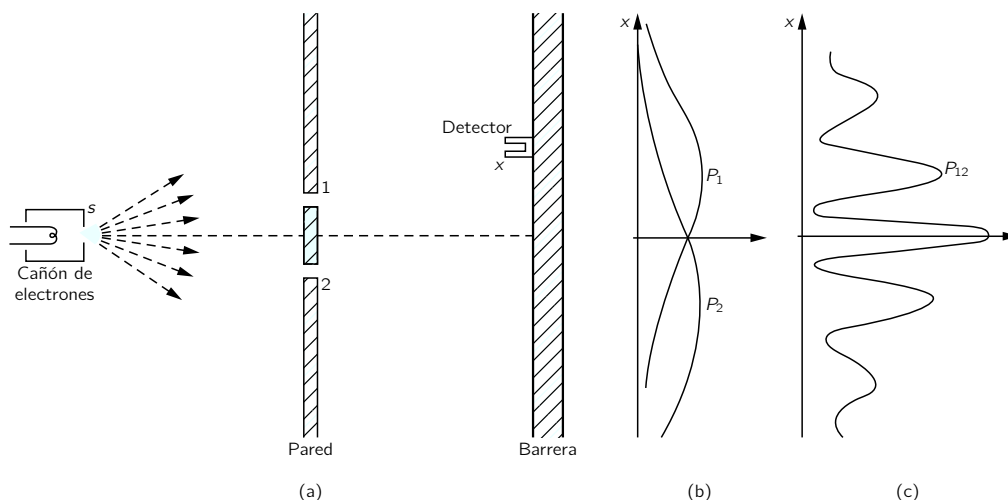


Figura 3-1. Experimento de interferencia con electrones.

las características de diferentes fenómenos. Aquí trataremos de encontrar un término medio entre los dos extremos.

Comenzaremos en este capítulo tratando algunas ideas cuánticas generales. Algunos de los argumentos serán bastante precisos, otros sólo parcialmente. Será difícil decirles, a medida que avancemos, cuál es cuál, pero para cuando hayan terminado el resto del libro, entenderán qué partes son consistentes y qué partes se explicaron sólo vagamente. Los capítulos que siguen a éste no serán tan imprecisos. En realidad, una de las razones por las que hemos tratado cuidadosamente de ser precisos en los capítulos siguientes, es para poder mostrarles uno de los aspectos más bellos de la mecánica cuántica —cuánto se puede deducir a partir de tan poco—.

Comenzamos discutiendo de nuevo la superposición de las *amplitudes de probabilidad*. Como ejemplo, nos referiremos al experimento descrito en el capítulo 1, el cual mostramos nuevamente en la figura 3-1. Hay una fuente s de partículas, digamos electrones; después hay una pared con dos rejillas en ella; detrás de la pared, hay un detector colocado en cierta posición x . Nos preguntamos por la probabilidad para que una partícula sea encontrada en x . Nuestro *primer principio general* en la mecánica cuántica, es que la *probabilidad* para que una partícula llegue a x , cuando sale de la fuente s , se puede representar cuantitativamente por el cuadrado de la norma de un número complejo, llamado *amplitud de probabilidad* —en este caso, la “amplitud para que una partícula de s llegue a x ”—. Usaremos dichas amplitudes con tanta frecuencia, que vamos a utilizar una notación abreviada —inventada por Dirac y usada generalmente en mecánica cuántica— para representar esta idea. Escribiremos la amplitud de probabilidad de esta manera:

$$\langle \text{Partícula que llega a } x \mid \text{partícula que sale de } s \rangle. \quad (3.1)$$

En otras palabras, los dos corchetes $\langle \rangle$ son un signo equivalente a “la amplitud para que”; la expresión a la *derecha* de la línea vertical siempre da la condición *inicial*, y la otra en la izquierda, la condición *final*. Algunas veces será conveniente abreviar aún más, y describir las condiciones inicial y final sólo con letras. Por ejemplo, podremos escribir, en alguna ocasión, la amplitud (3.1) así:

$$\langle x \mid s \rangle. \quad (3.2)$$

Queremos hacer notar que tal amplitud es, por supuesto, sólo un número —un número *complejo*—.

Ya hemos visto, en el capítulo 1, que, cuando existen dos caminos para que la partícula llegue al detector, la probabilidad resultante no es la suma de las

dos probabilidades, sino que se debe escribir como el cuadrado de la norma de la suma de dos amplitudes. Teníamos que la probabilidad para que un electrón llegue al detector, cuando ambos caminos están abiertos, es:

$$P_{12} = |\phi_1 + \phi_2|^2. \quad (3.3)$$

Ahora queremos escribir este resultado en términos de nuestra nueva notación. Antes, sin embargo, queremos establecer nuestro *segundo principio general* de la mecánica cuántica: cuando una partícula puede llegar a un estado dado por dos caminos posibles, la amplitud total para el proceso es la *suma de las amplitudes* para los dos caminos considerados por separado. En nuestra nueva notación escribimos que:

$$\langle x | s \rangle_{\text{ambos agujeros abiertos}} = \langle x | s \rangle_{\text{a través de 1}} + \langle x | s \rangle_{\text{a través de 2}}. \quad (3.4)$$

Además, vamos a suponer que los agujeros 1 y 2 son lo suficientemente pequeños, para que cuando digamos que un electrón pasa por un agujero, no tengamos que discutir por qué parte de él. Podríamos, naturalmente, dividir cada agujero en partes, con una cierta amplitud para que el electrón pase por la parte superior del agujero, o por la parte inferior, etc. Supondremos que el agujero es lo suficientemente pequeño, por lo que no tendremos que preocuparnos de este detalle. Esto es parte de la falta de detalle de la que hablamos; se podría tratar con más precisión, pero no queremos hacerlo a esta altura.

En este momento queremos escribir con más detalle lo que podemos decir acerca de la amplitud para el proceso, en el cual el electrón llega al detector en x , pasando por el agujero 1. Podemos hacer esto usando nuestro *tercer principio general*: cuando una partícula sigue un camino determinado, la amplitud para este camino se puede escribir como el *producto* de la *amplitud* para ir parte del camino por la *amplitud* para el resto del camino. Para el arreglo en la figura 3-1, la amplitud para ir desde s hasta 1 por el agujero 1, es igual a la amplitud para ir desde s hasta 1, multiplicada por la amplitud para ir desde 1 hasta x .

$$\langle x | s \rangle_{\text{vía 1}} = \langle x | 1 \rangle \langle 1 | s \rangle. \quad (3.5)$$

De nuevo, este resultado no es completamente preciso. También tendríamos que incluir un factor debido a la amplitud para que el electrón atraviese el agujero 1; pero en este caso es un agujero simple, y diremos que ese factor es uno.

Notarán que la ecuación (3.5) parece estar escrita en orden inverso. Se debe leer de derecha a izquierda: el electrón va de s a 1, y luego de 1 a x . En resumen, si los eventos ocurren en sucesión —es decir, si ustedes pueden analizar uno de los caminos de la partícula diciendo que hace esto, luego eso y luego aquello— la amplitud resultante para ese camino se calcula multiplicando en sucesión la amplitud para cada uno de los pasos sucesivos. Usando esta ley podemos volver a escribir la ecuación (3.4) así:

$$\langle x | s \rangle_{\text{ambos}} = \langle x | 1 \rangle \langle 1 | s \rangle + \langle x | 2 \rangle \langle 2 | s \rangle.$$

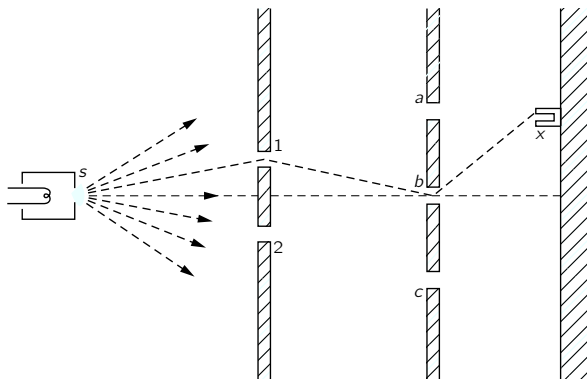


Figura 3-2. Un experimento de interferencia más complicado.

Ahora, queremos mostrar que, usando solamente estos principios, podemos calcular un problema mucho más complicado, como el mostrado en la figura 3-2. Aquí tenemos dos paredes, una con dos agujeros, 1 y 2, y otra que tiene tres agujeros, a , b y c . Detrás de la segunda pared tenemos un detector en x , y queremos conocer la amplitud para que una partícula llegue a él. Bien, un modo para encontrarla es calculando la superposición, o interferencia, de las ondas que pasan; pero también pueden hacerlo diciendo que existen seis posibles caminos, y superponiendo una amplitud por cada uno. El electrón puede atravesar el agujero 1, luego el a y llegar a x ; o podría atravesar el agujero 1, luego el b y llegar a x ; y continuar de este modo. De acuerdo con nuestro segundo principio, las amplitudes para diferentes caminos se suman, por lo que podríamos escribir la amplitud para ir desde s hasta x como una suma de seis amplitudes separadas. De otra manera, usando el tercer principio, cada una de estas amplitudes separadas se puede escribir como el producto de tres amplitudes. Por ejemplo, una de ellas es la amplitud para ir desde s a 1, multiplicada por la amplitud para ir de 1 a a , multiplicada por la amplitud para ir de a a x . Usando nuestra notación abreviada, podemos escribir la amplitud completa para ir desde s hasta x como

$$\langle x | s \rangle = \langle x | a \rangle \langle a | 1 \rangle \langle 1 | s \rangle + \langle x | b \rangle \langle b | 1 \rangle \langle 1 | s \rangle + \cdots + \langle x | c \rangle \langle c | 2 \rangle \langle 2 | s \rangle.$$

Nos ahorraremos la escritura usando la notación de suma

$$\langle x | s \rangle = \sum_{\substack{i=1,2 \\ \alpha=a,b,c}} \langle x | \alpha \rangle \langle \alpha | i \rangle \langle i | s \rangle. \quad (3.6)$$

Para poder hacer cálculos utilizando estos métodos, es necesario, naturalmente, conocer la amplitud para ir desde un lugar hasta otro. Daremos una idea aproximada de una amplitud típica. No consideramos ciertos aspectos, como la polarización de la luz o el espín del electrón, pero, aparte de estos aspectos, es bastante precisa. La damos para que ustedes puedan resolver problemas en los que haya diversas combinaciones de rejillas. Suponemos que una partícula con una determinada energía se mueve en el espacio vacío, desde una posición $\mathbf{r}_1$ hasta una posición $\mathbf{r}_2$. En otras palabras, es una partícula libre sobre la que no actúan fuerzas. Con excepción de un factor numérico enfrente, la amplitud para ir de $\mathbf{r}_1$ a $\mathbf{r}_2$ es

$$\langle \mathbf{r}_2 | \mathbf{r}_1 \rangle = \frac{e^{i\mathbf{p} \cdot \mathbf{r}_{12}/\hbar}}{r_{12}}, \quad (3.7)$$

donde $\mathbf{r}_{12} = \mathbf{r}_2 - \mathbf{r}_1$, y $\mathbf{p}$ es el momentum, el cual está relacionado con la energía E mediante la ecuación relativista:

$$p^2 c^2 = E^2 - (m_0 c^2)^2,$$

o mediante la ecuación no relativista:

$$\frac{p^2}{2m} = \text{Energía cinética}.$$

La ecuación (3.7) nos dice, en efecto, que la partícula tiene propiedades parecidas a las de las ondas, la amplitud se propaga como una onda, con un número de onda igual al momentum dividido entre $\hbar$.

En el caso más general, la amplitud y su correspondiente probabilidad dependerán también del tiempo. Para la mayoría de estas discusiones iniciales, supondremos que la fuente siempre emite partículas con una energía determinada, por lo que no tendremos que preocuparnos por el tiempo. Pero podríamos, en el caso general, estar interesados en algunas otras cuestiones. Supongan que una partícula se libera en un cierto lugar $\mathbf{P}$, en un cierto momento, y que quieren conocer la amplitud para que llegue a algún lugar, digamos $\mathbf{r}$, en un cierto momento después. Esto se podría representar simbólicamente como la amplitud $\langle \mathbf{r}, t = t_1 | \mathbf{P}, t = 0 \rangle$. Claramente, esto dependerá de $\mathbf{r}$ y t . Obtendrán diferentes resultados si colocan el detector en distintos lugares y miden en diferentes tiempos.

Esta función de $\mathbf{r}$ y t satisface, en general, una ecuación diferencial que es una ecuación de onda. Por ejemplo, en un caso no relativista, es la ecuación de Schrödinger. Tenemos, entonces, una ecuación de onda análoga a la ecuación para ondas electromagnéticas u ondas acústicas en un gas. No obstante, hay que notar que la función de onda que satisface la ecuación no es como una onda real en el espacio; no se le puede asociar ninguna clase de realidad a esta onda, como se le da a una onda acústica.

Aunque se pueda pensar en términos de “ondas-partículas” cuando se trata con una sola partícula, no es una buena idea, ya que si hay, por ejemplo, dos partículas, la amplitud para encontrar una en $\mathbf{r}_1$ y la otra en $\mathbf{r}_2$ no es una onda simple en un espacio tridimensional, sino que depende de las *seis* variables espaciales $\mathbf{r}_1$ y $\mathbf{r}_2$. Si estamos tratando, por ejemplo, con dos (o más) partículas, necesitaremos el siguiente principio adicional: dando por sentado que las dos partículas no interactúan, la amplitud para que una partícula haga una cosa *y* la otra partícula haga otra cosa, es el producto de las dos amplitudes para que las dos partículas hagan ambas cosas de manera separada. Por ejemplo, si $\langle a | s_1 \rangle$ es la amplitud para que la partícula 1 vaya de s_1 a a , y $\langle b | s_2 \rangle$ es la amplitud para que la partícula 2 vaya desde s_2 a b , la amplitud para que ambas cosas sucedan a la vez es

$$\langle a | s_1 \rangle \langle b | s_2 \rangle.$$

Hay otro punto que se debe resaltar. Supongamos que no sabemos de dónde llegan las partículas de la figura 3-2 antes de llegar a los agujeros 1 y 2 de la primera pared. Aún así podemos realizar una predicción de lo que sucederá después de la pared (por ejemplo, la amplitud para llegar a x), siempre y cuando tengamos dos números: la amplitud para haber llegado a 1 y la amplitud para haber llegado a 2. En otras palabras, ya que las amplitudes para eventos sucesivos se multiplican, como se muestra en la ecuación (3.6), todo lo que se necesita saber para continuar el análisis son dos números —en este caso particular $\langle 1 | s \rangle$ y $\langle 2 | s \rangle$ —. Estos dos números complejos nos bastan para predecir todo el futuro. Esto es lo que hace a la mecánica cuántica fácil. Resulta que, en los capítulos posteriores, eso es lo que haremos cuando especifiquemos una condición inicial en términos de dos (o más) números. Por supuesto, estos números dependerán del lugar donde esté colocada la fuente y, posiblemente, de otros detalles del aparato, pero dados los dos números, no necesitaremos saber más sobre esos detalles.

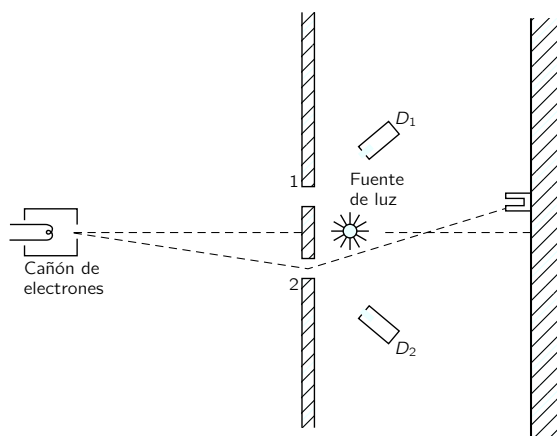


Figura 3-3. Un experimento para determinar por cuál agujero pasa el electrón.

3-2 El patrón de interferencia de dos rejillas

Ahora nos gustaría considerar un asunto que discutimos con algún detalle en el capítulo 1. Esta vez lo haremos con toda la gloria del concepto de amplitud para mostrarles cómo funciona. Tomamos el mismo experimento mostrado en la figura 3-1, pero ahora agregamos una fuente de luz colocada al otro lado de los agujeros, como se ve en la figura 3-3. En el capítulo 1 descubrimos el interesante resultado que sigue. Si observábamos detrás de la rejilla 1 y veíamos que ahí se dispersaba un fotón, entonces la distribución obtenida para los electrones en x ,

en coincidencia con estos fotones, era la misma como cuando la rejilla 2 estaba cerrada. La distribución total para los electrones que se habían “visto”, ya sea en la rejilla 1 o en la rejilla 2, era la suma de las distribuciones separadas, y era completamente diferente de la distribución obtenida con la luz apagada. Cuando usábamos luz de longitud de onda suficientemente pequeña esto era verdad. Si la longitud de onda se hacía más grande, de tal manera que no estuviéramos seguros por cuál agujero se había producido la dispersión, la distribución se parecía más a la obtenida con la luz apagada.

Examinemos qué sucede al usar nuestra nueva notación y los principios para combinar amplitudes. Para simplificar la escritura, supondremos de nuevo que ϕ_1 es la amplitud para que el electrón llegue a x por el agujero 1; esto es

$$\phi_1 = \langle x | 1 \rangle \langle 1 | s \rangle.$$

Del mismo modo, supondremos que ϕ_2 representa a la amplitud para que el electrón llegue al detector a través del agujero 2:

$$\phi_2 = \langle x | 2 \rangle \langle 2 | s \rangle.$$

Éstas son las amplitudes para ir a través de los dos agujeros y llegar a x si no hay luz. Ahora bien, si hay luz, nos hacemos la siguiente pregunta: ¿cuál es la amplitud para el proceso en el cual un electrón sale de s y se libera un fotón en la fuente de luz L , que termina con el electrón en x y viéndose un fotón en la parte posterior de la rejilla 1? Ahora supongamos que observamos al fotón detrás de la rejilla 1 por medio de un detector D_1 , como se muestra en la figura 3-3, y que usamos un detector similar, D_2 , para contar los fotones dispersados detrás de la rejilla 2. Habrá una amplitud para que un fotón llegue a D_1 y un electrón a x , y también otra amplitud para que un fotón llegue a D_2 y un electrón a x . Intentemos calcularlas.

Aunque no tenemos las fórmulas matemáticas correctas para todos los factores que intervienen en este cálculo, verán la esencia en la siguiente discusión. En primer lugar, tenemos la amplitud $\langle 1 | s \rangle$ para que un electrón vaya desde la fuente hasta el agujero 1. Entonces, podemos suponer que existe una cierta amplitud para que, mientras el electrón está en el agujero 1, éste disperse un fotón hacia el detector D_1 . Representamos a esta amplitud por a . También tenemos la amplitud $\langle x | 1 \rangle$ para que el electrón vaya del agujero 1 al detector de electrones en x . La amplitud para que el electrón vaya desde s hasta x a través del agujero 1 *y además* disperse un fotón hacia D_1 , es entonces

$$\langle x | 1 \rangle a \langle 1 | s \rangle.$$

O, en nuestra notación anterior, es simplemente $a\phi_1$.

También hay una cierta amplitud para que un electrón que atraviesa el agujero 2 disperse un fotón hacia el contador D_1 . Ustedes dirán: “Eso es imposible; ¿cómo puede dispersar hacia el contador D_1 , si éste solamente está viendo hacia el agujero 1?” Si la longitud de onda es lo suficientemente larga, existen efectos de difracción, y en realidad es posible. Si el aparato está bien construido, y usamos fotones de longitud de onda corta, entonces la amplitud para que un fotón sea dispersado hacia el detector 1, por un electrón que atraviesa 2, es muy pequeña. Sin embargo, para continuar la discusión de manera general, queremos tomar en cuenta que siempre existe dicha amplitud, que llamaremos b . Entonces, la amplitud para que un electrón atraviese la rejilla 2 *y también* disperse un fotón hacia D_1 es:

$$\langle x | 2 \rangle b \langle 2 | s \rangle = b\phi_2.$$

La amplitud para encontrar al electrón en x y al fotón en D_1 es la suma de los dos términos, uno para cada camino posible del electrón. Cada término está, a su vez, formado por dos factores: primero, que el electrón pase a través de un agujero y, segundo, que el fotón sea dispersado por tal electrón hacia el detector 1; tenemos:

$$\left\langle \begin{array}{l} \text{electrón en } x \\ \text{fotón en } D_1 \end{array} \middle| \begin{array}{l} \text{electrón de } s \\ \text{fotón de } L \end{array} \right\rangle = a\phi_1 + b\phi_2. \quad (3.8)$$

Podemos obtener una expresión similar cuando encontramos al fotón en el detector D_2 . Si suponemos, para simplificar, que el sistema es simétrico, entonces a también es la amplitud que corresponde a un fotón en D_2 cuando un electrón pasa por el agujero 2, y b es la amplitud para un fotón en el detector D_2 cuando el electrón atraviesa el agujero 1. La amplitud total correspondiente para un fotón en D_2 y un electrón en x es:

$$\left\langle \begin{array}{c} \text{electrón en } x \\ \text{fotón en } D_2 \end{array} \middle| \begin{array}{c} \text{electrón de } s \\ \text{fotón de } L \end{array} \right\rangle = a\phi_2 + b\phi_1. \quad (3.9)$$

Ya hemos terminado. Podemos calcular fácilmente la probabilidad para diversas situaciones. Supongamos que queremos saber con qué probabilidad contamos un fotón en D_1 y un electrón en x . Esto será el cuadrado de la norma de la amplitud dada en la ecuación (3.8), es decir, $|a\phi_1 + b\phi_2|^2$. Examinemos más cuidadosamente esta expresión. Primero, si b es cero —que sería el modo en el que nos gustaría diseñar al aparato— entonces la respuesta es simplemente $|\phi_1|^2$, disminuida por el factor $|a|^2$. Ésta es la distribución de probabilidad que obtendrían si hubiese sólo un agujero —como se muestra en el gráfico de la figura 3-4(a)—. Por otra parte, si la longitud de onda es muy grande, la dispersión detrás del agujero 2 hacia D_1 sería casi la misma que para el agujero 1. Aunque puede haber algunas fases en a y en b , podemos considerar el caso simple en el que las dos fases son iguales. Si a es prácticamente igual a b , entonces la probabilidad total es $|\phi_1 + \phi_2|^2$ multiplicada por $|a|^2$, ya que el factor común a se puede sacar. Ésta, sin embargo, es exactamente la distribución de probabilidad que habríamos obtenido sin los fotones. Por lo tanto, en el caso de que la longitud de onda sea muy grande —y la detección del fotón es ineficiente— regresan a la curva original de distribución que muestra efectos de interferencia, como se ve en la figura 3-4(b). En el caso cuando la detección es parcialmente eficiente, existe interferencia entre gran parte de ϕ_1 y una pequeña parte de ϕ_2 , y obtendrán una distribución intermedia como la mostrada en la figura 3-4(c). No hay necesidad de decir que, si buscamos detectar simultáneamente fotones en D_2 y electrones en x , obtendremos resultados semejantes. Si recuerdan la discusión del capítulo 1, verán que estos resultados nos dan una descripción cuantitativa de lo que se describió allá.

Ahora queremos remarcar un punto importante, para que eviten un error común. Supongan que sólo quieren la amplitud para que el electrón llegue a x , *sin importarles* si el fotón se contó en D_1 o en D_2 . ¿Deberían sumar las amplitudes dadas en las ecuaciones (3.8) y (3.9)? ¡No! *Nunca deben sumar amplitudes para estados finales diferentes y distintos*. Una vez que el fotón ha sido aceptado por cualquiera de los contadores de fotones, siempre podemos determinar qué alternativa ocurrió si así lo queremos, sin perturbar en lo absoluto al sistema. Cada alternativa tiene una probabilidad completamente independiente de la otra. Por repetir, no sumen amplitudes para condiciones *finales* diferentes, donde por “final” entendemos el momento en que se desea la *probabilidad* —esto es, cuando el experimento ha “terminado”—. Sí suman las amplitudes para diferentes alternativas *indistinguibles* dentro del experimento, antes de que termine el proceso completo. Al final del proceso pueden decir que ustedes “no quieren observar al fotón”. Ése es su problema, pero aun así no suman las amplitudes. La naturaleza no sabe qué es lo que están observando, y se comporta del modo que se va a comportar, ya sea que ustedes se preocupen o no de anotar los resultados. De modo que aquí no debemos sumar las amplitudes. Primero calculamos las normas de las amplitudes para todos los posibles eventos finales diferentes, y luego las sumamos. El resultado correcto para un electrón en x y un fotón ya sea en D_1 o en D_2 es

$$\begin{aligned} & \left| \left\langle \begin{array}{c} e \text{ en } x \\ f \text{ en } D_1 \end{array} \middle| \begin{array}{c} e \text{ de } s \\ f \text{ de } L \end{array} \right\rangle \right|^2 + \left| \left\langle \begin{array}{c} e \text{ en } x \\ f \text{ en } D_2 \end{array} \middle| \begin{array}{c} e \text{ de } s \\ f \text{ de } L \end{array} \right\rangle \right|^2 \\ & = |a\phi_1 + b\phi_2|^2 + |a\phi_2 + b\phi_1|^2. \end{aligned} \quad (3.10)$$

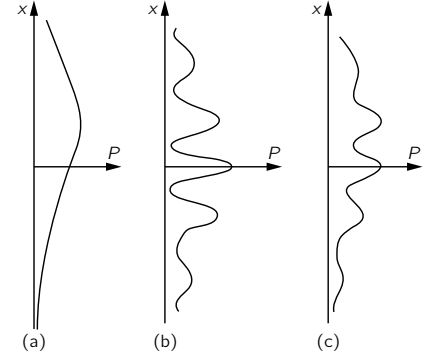


Figura 3-4. La probabilidad para contar un electrón en x al mismo tiempo con un fotón en D en el experimento de la figura 3-3: (a) para $b = 0$; (b) para $b = a$; (c) para $0 < b < a$.

3-3 Dispersión en un cristal

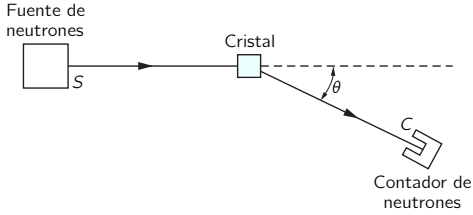


Figura 3-5. Medición de la dispersión de neutrones por un cristal.

Nuestro próximo ejemplo es un fenómeno en el cual tenemos que analizar la interferencia en las amplitudes de probabilidad de un modo un tanto cuidadoso. Nos referimos al proceso de la dispersión de neutrones en un cristal. Si tenemos un cristal con varios átomos con sus núcleos en el centro, ordenados en un arreglo periódico, y un haz de neutrones que llega de muy lejos, entonces podemos designar a los diversos núcleos del cristal con un índice i , donde i recorre los números enteros $1, 2, 3, \dots, N$, con N igual al número total de átomos. El problema consiste en calcular la probabilidad para que un neutrón entre en un contador con el arreglo mostrada de la figura 3-5. Para cualquier átomo particular i , la amplitud para que el neutrón llegue al contador C es la amplitud para que el neutrón llegue desde la fuente S hasta el núcleo i , multiplicada por la amplitud a para que sea dispersado ahí, multiplicada por la amplitud para que llegue desde i hasta el contador C . Escribámoslo:

$$\langle \text{neutrón en } C \mid \text{neutrón de } S \rangle_{\text{vía } i} = \langle C \mid i \rangle a \langle i \mid S \rangle. \quad (3.11)$$

Al escribir esta ecuación hemos supuesto que la amplitud para la dispersión, a , es la misma para todos los átomos. Tenemos aquí un gran número de caminos aparentemente indistinguibles. Éstos son indistinguibles porque un neutrón de baja energía es dispersado por un núcleo sin sacar el átomo de su lugar en el cristal —no deja “señal” de la dispersión—. De acuerdo con nuestra discusión anterior, la amplitud total para un neutrón en C implica una suma de la ecuación (3.11) sobre todos los átomos:

$$\langle \text{neutrón en } C \mid \text{neutrón de } S \rangle = \sum_{i=1}^N \langle C \mid i \rangle a \langle i \mid S \rangle. \quad (3.12)$$

Como estamos sumando amplitudes para la dispersión desde átomos con diferentes posiciones en el espacio, las amplitudes tendrán fases diferentes, produciendo el patrón de interferencia característico, el cual ya hemos analizado en el caso de la dispersión de la luz por una rejilla.

La intensidad de los neutrones como una función del ángulo, en un experimento como éste, con frecuencia muestra grandes variaciones, con picos muy agudos de interferencia y casi nada entre ellos —como se muestra en la figura 3-6(a)—. Sin embargo, para cierta clase de cristales no ocurre así, y existe —junto con los picos de interferencia antes mencionados— un fondo general de dispersión en todas direcciones. Debemos tratar de entender las razones, aparentemente misteriosas, para esto. Bien, no hemos considerado una propiedad importante del neutrón. Tiene espín un medio, por lo que hay dos condiciones en las que puede estar: ya sea con el espín “hacia arriba” (digamos perpendicular a la página en la figura 3-5) o con el espín “hacia abajo”. Si los núcleos del cristal no tienen espín, el espín del neutrón no produce ningún efecto. Pero cuando los núcleos del cristal también tienen espín, digamos con espín un medio, observarán el fondo borroso de dispersión mencionado antes. La explicación es la siguiente.

Si el neutrón tiene una dirección de espín y el núcleo atómico tiene el mismo espín, entonces no puede ocurrir un cambio de espín en el proceso de dispersión. Si, en cambio, el neutrón y el núcleo atómico tienen espines opuestos, entonces la dispersión puede presentarse mediante dos procesos, uno en el cual los espines no cambian, y otro en el cual las direcciones de los espines se intercambian. Esta regla para que no haya cambio neto en la suma de los espines es análoga a nuestra ley clásica de la conservación del momentum angular. Podemos comenzar a comprender el fenómeno si suponemos que todos los núcleos que dispersan están arreglados con los espines en la misma dirección. Un neutrón con el mismo espín se dispersará con la distribución de interferencia bien definida que esperábamos. ¿Qué ocurre a uno con espín opuesto? Si se dispersa sin cambio en el espín, entonces no cambia nada del caso anterior; pero si los dos espines se intercambian en la dispersión, podríamos, en principio, encontrar cuál núcleo ha efectuado la dispersión, puesto que sería el único con el espín en dirección opuesta. Bien, si

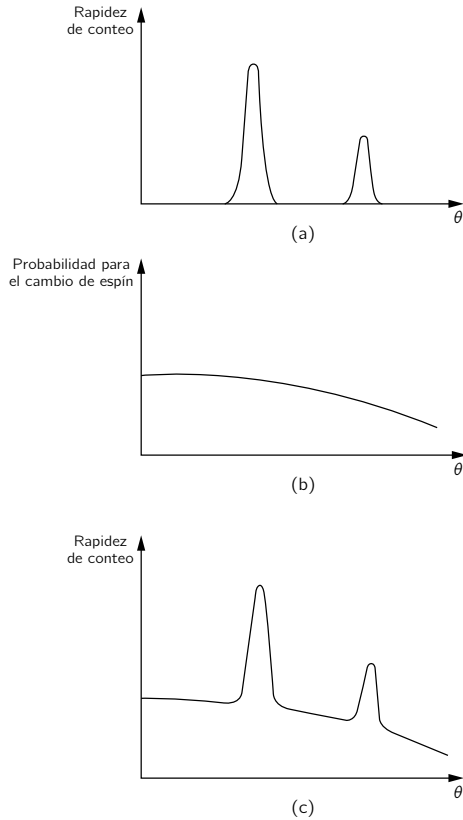


Figura 3-6. La rapidez de conteo de neutrones en función del ángulo: (a) para núcleos con espín cero; (b) la probabilidad para la dispersión con cambio de espín; (c) la rapidez de conteo observada para un núcleo con espín un medio.

podemos decir cuál átomo efectuó la dispersión, ¿qué tienen que ver los otros átomos con él? Nada, naturalmente. La dispersión es exactamente la misma que la de un solo átomo.

Para incluir este efecto, la formulación matemática de la ecuación (3.12) debe ser modificada, ya que no describimos completamente los estados en aquel análisis. Comencemos con todos los neutrones de la fuente con los espines hacia arriba, y todos los núcleos del cristal con los espines hacia abajo. En primer lugar, queremos una amplitud para que, en el contador, el espín del neutrón esté hacia arriba, *y además* que todos los espines del cristal estén hacia abajo. Esto no difiere en nada de nuestra discusión anterior. Llamemos a a la amplitud para la dispersión sin cambios o sin espín. La amplitud para que el átomo i disperse es, naturalmente,

$$\langle C_{\text{arriba}}, \text{cristal hacia abajo} | S_{\text{arriba}}, \text{cristal hacia abajo} \rangle = \langle C | i \rangle a \langle i | S \rangle.$$

Como todos los espines atómicos aún están hacia abajo, las diversas alternativas (diferentes valores de i) no se pueden distinguir. No hay una manera clara de decir cuál átomo realizó la dispersión. En este proceso, todas las amplitudes interfieren.

Tenemos otro caso, no obstante, en el cual el espín del neutrón en el detector está hacia abajo, aunque salió de S con su espín hacia arriba. En el cristal, uno de los espines debe haber cambiado a la dirección hacia arriba —digamos, el del k -ésimo átomo—. Supondremos que existe la misma amplitud para la dispersión con cambio en el espín para cada átomo, digamos b . (En un cristal real, existe la desagradable posibilidad de que el cambio en el espín se mueva a otro átomo, pero tomemos el caso de un cristal para el cual esta posibilidad es muy pequeña.) La amplitud para la dispersión es entonces:

$$\langle C_{\text{abajo}}, \text{núcleo } k \text{ hacia arriba} | S_{\text{arriba}}, \text{cristal hacia abajo} \rangle = \langle C | k \rangle b \langle k | S \rangle. \quad (3.13)$$

Si preguntamos por la probabilidad para encontrar al espín del neutrón hacia abajo y al espín del k -ésimo átomo hacia arriba, es igual al cuadrado de la norma de esta amplitud, la cual es simplemente $|b|^2$ por $|\langle C | k \rangle \langle k | S \rangle|^2$. El segundo factor es casi independiente de la posición en el cristal, y todas las fases han desaparecido al tomar el cuadrado de la norma. La probabilidad para la dispersión en *cualquier núcleo* del cristal con cambio en el espín es ahora

$$|b|^2 \sum_{k=1}^N |\langle C | k \rangle \langle k | S \rangle|^2,$$

que tendrá una distribución suave como en la figura 3-6(b).

Pueden argumentar: “No nos importa cuál átomo tiene el espín hacia arriba”. Quizá no les importe, pero la naturaleza sabe; y la probabilidad es, de hecho, la que dimos arriba —no hay ninguna interferencia—. Por otra parte, si preguntamos por la probabilidad para que el espín esté hacia arriba en el detector, y que todos los átomos aún tengan espín hacia abajo, entonces debemos tomar el cuadrado de la norma de

$$\sum_{i=1}^N \langle C | i \rangle a \langle i | S \rangle.$$

Puesto que los términos en esta suma tienen fases, sí interfieren, y obtenemos un patrón de interferencia bien definido. Si llevamos a cabo un experimento en el que no observamos el espín del neutrón en el detector, entonces ambos eventos pueden ocurrir; y las probabilidades separadas se suman. La probabilidad total (o la rapidez de conteo) en función del ángulo se parece a la gráfica de la figura 3-6(c).

Veamos el aspecto físico de este experimento. Si ustedes pueden, *en principio*, distinguir los diversos estados *finales* (aunque no se preocupen en hacerlo), la probabilidad total final se obtiene calculando la *probabilidad* para cada estado (no la amplitud) y luego sumándolas. Si *no pueden* distinguir los estados finales, *aun*

en principio, entonces las amplitudes de probabilidad se deben sumar antes de tomar el cuadrado de la norma para encontrar la probabilidad real. Lo que deben notar particularmente es que, si tratan de representar al neutrón sólo por como una onda, obtendrán la misma distribución para la dispersión de un neutrón con espín hacia abajo que con espín hacia arriba. Tendrían que decir que la “onda” proviene de todos los diferentes átomos, y que interfiere precisamente como si tuviera espín hacia arriba, y con la misma longitud de onda. Pero nosotros sabemos que no es así como sucede. Como dijimos antes, debemos tener cuidado de no atribuir demasiada realidad a las ondas en el espacio. Son muy útiles para ciertos problemas, pero no para todos.

3-4 Partículas idénticas

El próximo experimento que describiremos muestra una de las más bellas consecuencias de la mecánica cuántica. De nuevo implica una situación física en la que algo puede ocurrir de dos maneras *indistinguibles*, de tal forma que hay interferencia de amplitudes —como ocurre *siempre* en tales circunstancias—. Vamos a discutir la dispersión, a energía relativamente baja, de núcleos por otros núcleos. Empecemos considerando partículas α (las cuales son núcleos de helio, como saben) que bombardean, por ejemplo, oxígeno. Para que sea más fácil analizar la reacción, nos enfocaremos en el centro de masa del sistema, en el cual el núcleo de oxígeno y la partícula α tienen sus velocidades en direcciones opuestas antes de la colisión, y de nuevo en direcciones exactamente opuestas después de la colisión. Véase la figura 3-7(a). (Las magnitudes de las velocidades son, naturalmente, diferentes, puesto que las masas son distintas.) También supondremos que hay conservación de la energía, y que la energía de colisión es lo bastante pequeña como para que ninguna partícula se rompa o quede en un estado excitado. La razón para que las dos partículas se aparten mutuamente es, naturalmente, que cada partícula tiene una carga positiva y, hablando de manera clásica, existe una repulsión eléctrica cuando están cerca. La dispersión tendrá lugar a diferentes ángulos y con distintas probabilidades, y queremos discutir algo sobre la dependencia angular de dichas dispersiones. (Es posible, desde luego, calcular esto clásicamente, y eso es una de las casualidades más notables de la mecánica cuántica, pues la respuesta a este problema es la misma que la clásica. Éste es un hecho interesante, ya que sólo sucede para fuerzas que siguen la ley del inverso del cuadrado de la distancia —como decíamos, es realmente una casualidad—.)

La probabilidad para la dispersión en diferentes direcciones se puede medir mediante un experimento como el mostrado en la figura 3-7(a). El contador de

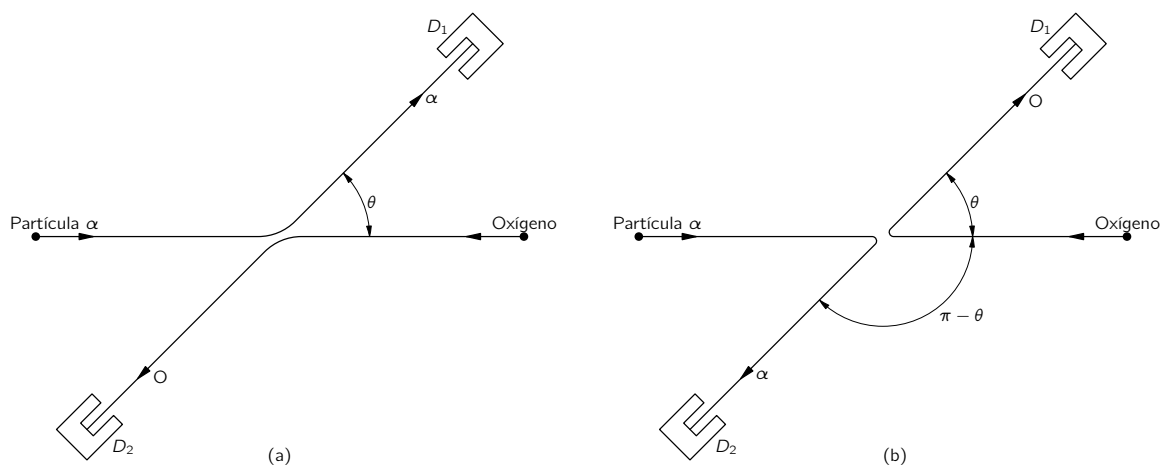


Figura 3-7. Dispersión de partículas α por núcleos de oxígeno, vista desde el centro de masa del sistema.

la posición 1 se podría diseñar para que sólo detectara partículas α ; el contador de la posición 2 se podría diseñar para detectar únicamente oxígeno —sólo como verificación—. (En el sistema de laboratorio los detectores no estarían opuestos; pero en el sistema CM lo están.) Nuestro experimento consiste en medir la probabilidad para la dispersión en diversas direcciones. Llamemos $f(\theta)$ a la amplitud para dispersar hacia los contadores cuando están en un ángulo θ ; entonces $|f(\theta)|^2$ será nuestra probabilidad determinada experimentalmente.

Ahora podríamos preparar otro experimento en el que nuestros contadores respondiesen *tanto* a la partícula α *como* al núcleo de oxígeno. Entonces, tenemos que ver lo que sucede cuando no nos preocupamos de distinguir cuáles partículas se están contando. Por supuesto, si vamos a obtener un oxígeno en la posición θ , debe haber una partícula α en el lado opuesto, en un ángulo $(\pi - \theta)$, como se muestra en la figura 3-7(b). Así, si $f(\theta)$ es la amplitud para la dispersión de la partícula α en un ángulo θ , entonces $f(\pi - \theta)$ es la amplitud para la dispersión del oxígeno en un ángulo θ .[†] Entonces, la probabilidad para que tengamos *alguna* partícula en el detector de la posición 1 es:

$$\text{Probabilidad para alguna partícula en } D_1 = |f(\theta)|^2 + |f(\pi - \theta)|^2. \quad (3.14)$$

Tengan presente que los dos estados son distinguibles en principio. Aunque en este experimento *no* los distinguimos, *podríamos hacerlo*. De acuerdo con la discusión anterior, entonces, tenemos que sumar las probabilidades, no las amplitudes.

El resultado de arriba es correcto para una variedad de núcleos que sirvan de objetivos —para partículas α con oxígeno, con carbono, con berilio, con hidrógeno—. *Pero es erróneo* para partículas α con partículas α . Porque el caso en que ambas partículas son exactamente iguales, los datos experimentales no concuerdan con la predicción de (3.14). Por ejemplo, la probabilidad para la dispersión a 90° es exactamente dos veces lo que predice la teoría anterior, y no tiene nada que ver con que las partículas sean núcleos de “helio”. Si el objetivo es He^3 , y los proyectiles son partículas α (He^4), entonces sí concuerdan. Sólo cuando el objetivo es He^4 —de manera que sus núcleos son idénticos a la partícula α que llega— la dispersión varía en una forma muy peculiar con el ángulo.

Quizá ya puedan ver la explicación. Hay dos formas para que la partícula α llegue al contador: dispersando la partícula α bombardeante en un ángulo θ , o dispersándola en un ángulo $(\pi - \theta)$. ¿Cómo podríamos decir si fue la partícula bombardeante o la partícula objetivo la que entró en el contador? La respuesta es que no podemos. En el caso de partículas α con partículas α , hay dos alternativas indistinguibles. Aquí, debemos permitir que las *amplitudes* para la probabilidad interfieran sumándose, y la probabilidad de encontrar una partícula α en el contador es el cuadrado de la suma:

$$\text{Probabilidad para una partícula } \alpha \text{ en } D_1 = |f(\theta) + f(\pi - \theta)|^2. \quad (3.15)$$

Éste es un resultado bastante diferente al de la ecuación (3.14). Podemos tomar un ángulo de $\pi/2$ como ejemplo, ya que es fácil de representar. Para $\theta = \pi/2$, obviamente obtenemos $f(\theta) = f(\pi - \theta)$, de tal forma que la probabilidad en la ecuación (3.15) se vuelve $|f(\pi/2) + f(\pi/2)|^2 = 4|f(\pi/2)|^2$.

Por otra parte, si no interfiriera, el resultado de la ecuación (3.14) solamente da $2|f(\pi/2)|^2$. Así que hay el doble de dispersión para 90° , como era de esperar. Naturalmente, para otros ángulos los resultados también serán diferentes. Y, por lo tanto, tienen el inusual resultado de que, cuando las partículas son idénticas, sucede algo nuevo que no ocurre cuando las partículas se pueden distinguir. En la descripción matemática tienen que sumar las amplitudes para

[†] En general, una dirección de dispersión se debería, naturalmente, describir mediante dos ángulos, el ángulo polar ϕ y el ángulo azimutal θ . Diríamos entonces que un núcleo de oxígeno situado en (θ, ϕ) , significa que la partícula α se encuentra en $(\pi - \theta, \phi + \pi)$. Sin embargo, para la dispersión de coulomb (y para muchos otros casos), la amplitud para la dispersión es independiente de ϕ . Así la amplitud para obtener un oxígeno en θ , es la misma que la amplitud para obtener una partícula α en $(\pi - \theta)$.

procesos alternativos, en los cuales las partículas sólo cambian sus papeles y hay interferencia.

Una cosa aún más asombrosa ocurre cuando realizamos el mismo experimento dispersando electrones con electrones, o protones con protones. ¡Ninguno de los resultados anteriores es entonces correcto! Para estas partículas, debemos usar una regla nueva, una regla muy peculiar, que es la siguiente: cuando tienen una situación en la cual la identidad del electrón que llega a un punto se cambia con la de otro, la nueva amplitud interfiere con la antigua con *fase opuesta*. Sí es una interferencia, pero con signo menos. En el caso de las partículas α , cuando ustedes intercambian a la partícula α que entra en el detector, las amplitudes interfieren con signo positivo. *En el caso de electrones, las amplitudes para el intercambio interfieren con un signo negativo*. Exceptuando un detalle que discutiremos luego, la ecuación correcta para los electrones, en un experimento como el mostrado en la figura 3-8, es

$$\text{Probabilidad para un e en } D_1 = |f(\theta) - f(\pi - \theta)|^2. \quad (3.16)$$

El enunciado anterior debe ser modificado, ya que no hemos considerado el espín del electrón (las partículas α no tienen espín). El espín del electrón se puede considerar que está “hacia arriba” o “hacia abajo” respecto al plano de la dispersión. Si la energía del experimento es bastante baja, las fuerzas magnéticas debidas a las corrientes serán pequeñas y el espín no se verá afectado. Supondremos que éste es el caso en el análisis actual, así que no habrá posibilidad para que los espines cambien durante la colisión. Cualquiera sea el espín que tenga el electrón, siempre lo lleva consigo. Ahora observen que hay muchas posibilidades. Las partículas que bombardean y las objetivo pueden tener ambos espines hacia arriba, ambos hacia abajo, o espines opuestos. Si ambos espines están hacia arriba, como en la figura 3-8 (o si ambos están hacia abajo), lo mismo sucederá para las partículas en los contadores, y la *amplitud* para el proceso es la *diferencia* entre las amplitudes para las dos posibilidades mostradas en la figura 3-8(a) y (b). La *probabilidad* para detectar un electrón en D_1 está dada entonces por la ecuación (3.16).

Supongan, sin embargo, que el espín de la partícula “que bombardea” está hacia arriba, y que el espín de la partícula “objetivo” está hacia abajo. El electrón que entra al contador 1 puede tener el espín hacia arriba o hacia abajo, y midiendo este espín podemos decir si proviene del haz que bombardea o del objetivo. Las dos posibilidades se muestran en la figura 3-9(a) y (b); en principio son distinguibles y, por lo tanto, no habrá interferencia —simplemente una suma de las dos probabilidades—. El mismo argumento funciona si los espines originales

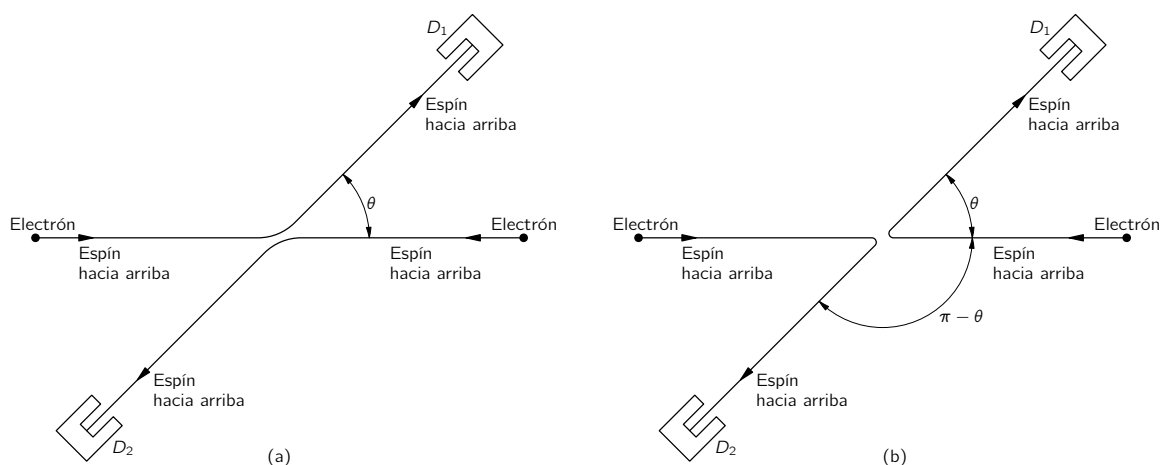


Figura 3-8. Dispersión de electrones con electrones. Si los electrones incidentes tienen espines paralelos, los procesos (a) y (b) son indistinguibles.

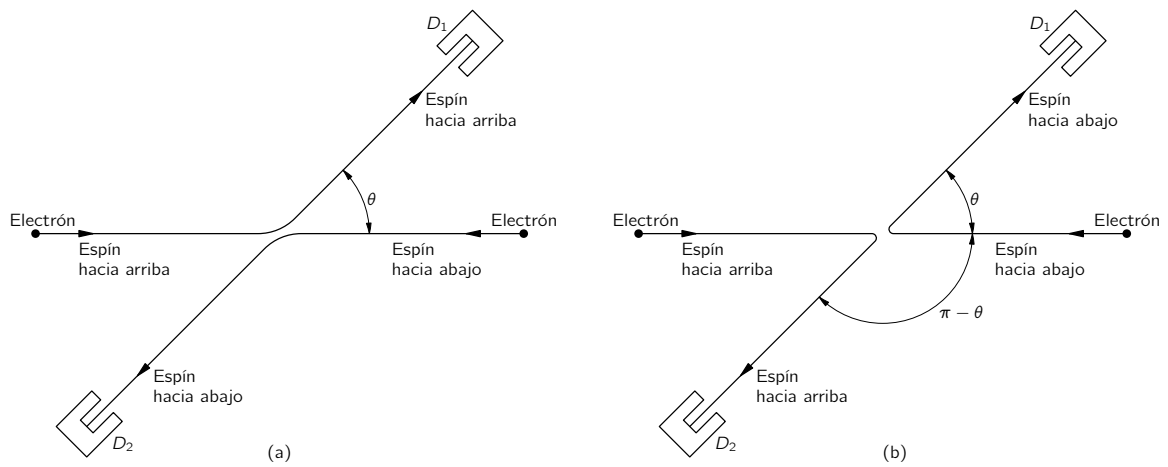


Figura 3-9. Dispersión de electrones con los espines antiparalelos.

están invertidos —es decir, si el espín de la izquierda está hacia abajo y el de la derecha hacia arriba—.

Ahora, si tomamos nuestros electrones al azar —como de un filamento de tungsteno, en el que los electrones están completamente sin polarizar— entonces la probabilidad es cincuenta-cincuenta para que algún electrón particular salga con espín hacia arriba o hacia abajo. Si no nos molestamos por medir el espín de los electrones en ningún punto en el experimento, tenemos lo que llamamos un experimento sin polarizar. Los resultados de este experimento se calculan mejor haciendo una lista de las diversas posibilidades, como hemos hecho en la tabla 3-1. Una *probabilidad* separada se calcula para cada alternativa distinguible. La probabilidad total es, entonces, la suma de todas las probabilidades por separado. Observen que, para haces no polarizados, el resultado para $\theta = \pi/2$ es la mitad del resultado clásico con partículas independientes. El comportamiento de las partículas idénticas tiene muchas consecuencias interesantes; las discutiremos con más detalle en el próximo capítulo.

Tabla 3-1

Dispersión de partículas no polarizadas de espín de un medio

Fracción de casos	Espín de la partícula 1	Espín de la partícula 2	Espín en D_1	Espín en D_2	Probabilidad
$\frac{1}{4}$	arriba	arriba	arriba	arriba	$ f(\theta) - f(\pi - \theta) ^2$
$\frac{1}{4}$	abajo	abajo	abajo	abajo	$ f(\theta) - f(\pi - \theta) ^2$
$\frac{1}{4}$	arriba	abajo	arriba	abajo	$ f(\theta) ^2$
			abajo	arriba	$ f(\pi - \theta) ^2$
$\frac{1}{4}$	abajo	arriba	arriba	abajo	$ f(\pi - \theta) ^2$
			abajo	arriba	$ f(\theta) ^2$
Probabilidad total = $\frac{1}{2} f(\theta) - f(\pi - \theta) ^2 + \frac{1}{2} f(\theta) ^2 + \frac{1}{2} f(\pi - \theta) ^2$					

Partículas idénticas

4-1 Bosones y fermiones

En el capítulo anterior, consideramos las reglas especiales para la interferencia que ocurre en procesos con dos partículas *idénticas*. Por *partículas idénticas* entendemos cosas tales como que los electrones no se pueden distinguir uno de otro. Si en un proceso intervienen dos partículas idénticas, el cambiar cual llega a un contador es una alternativa que no se puede distinguir y que —como en todos los casos con alternativas que no se pueden distinguir— interfiere con el caso original sin intercambiar. La amplitud de un evento es entonces la suma de las dos amplitudes que interfieren; pero, es interesante señalar que la interferencia en algunos casos se da con la *misma* fase y en otros con la fase *opuesta*.

Supongamos que tenemos una colisión entre dos partículas a y b , en la cual la partícula a se dispersa en la dirección 1 y la partícula b se dispersa en la dirección 2, como se indica en la figura 4-1(a). Llamemos $f(\theta)$ la amplitud para este proceso; entonces la probabilidad P_1 de observar tal evento es proporcional a $|f(\theta)|^2$. Por supuesto, también podría ocurrir que la partícula b se dispersara entrando en el contador 1 y la partícula a fuera al contador 2, como se muestra en la figura 4-1(b). Suponiendo que no hay direcciones especiales definidas por el espín o cosas por el estilo, la probabilidad P_2 para este proceso es simplemente $|f(\pi - \theta)|^2$, porque éste es equivalente al primer proceso con el contador 1 movido un ángulo $\pi - \theta$. Ustedes podrían pensar que la *amplitud* del segundo proceso es sólo $f(\pi - \theta)$. Pero no es necesariamente así, porque puede haber un factor de fase arbitrario. Es decir, la amplitud podría ser

$$e^{i\delta} f(\pi - \theta).$$

Esta amplitud también da una probabilidad P_2 igual a $|f(\pi - \theta)|^2$.

Ahora veamos lo que ocurre si a y b son partículas idénticas. En ese caso, los dos procesos diferentes mostrados en los dos diagramas de la figura 4-1 no se pueden distinguir. Hay una amplitud para que, *ya sea*, a o b vaya al contador 1, mientras que la otra vaya al contador 2. Esta amplitud es la suma de las amplitudes para los dos procesos mostrados en la figura 4-1. Si llamamos a la primera $f(\theta)$, entonces la segunda es $e^{i\delta} f(\pi - \theta)$, donde ahora el factor de

4-1 Bosones y fermiones

4-2 Estados con dos bosones

4-3 Estados con n bosones

4-4 Emisión y absorción de fotones

4-5 El espectro de cuerpo negro

4-6 El helio líquido

4-7 El principio de exclusión

Revisar

Capítulo 41, volumen I,

El movimiento browniano.

Capítulo 42, volumen I,

Aplicaciones de la teoría cinética.

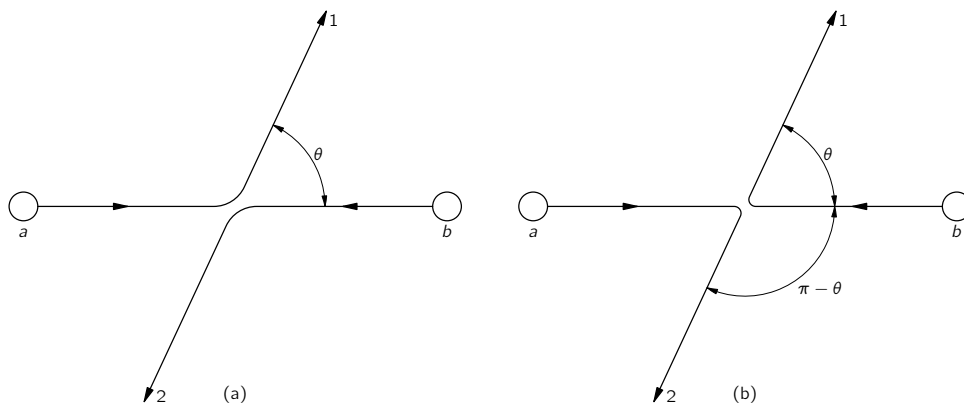


Figura 4-1. Dispersión de dos partículas idénticas, los procesos (a) y (b) son indistinguibles.

fase es muy importante porque vamos a sumar dos amplitudes. Supongamos que tenemos que multiplicar a la amplitud por un cierto factor de fase cuando intercambiamos los papeles de las dos partículas. Si las intercambiamos una vez más, deberíamos obtener el mismo factor de nuevo. Pero entonces hemos vuelto al primer proceso. El factor de fase aplicado dos veces debe regresarnos a donde empezamos —su cuadrado debe ser igual a 1—. Hay sólo dos posibilidades: $e^{i\delta}$ es igual a $+1$ o es igual a -1 . Ya sea que el caso intercambiado contribuya con el *mismo* signo, o contribuya con el signo *opuesto*. Ambos casos existen en la naturaleza, cada uno para una clase diferente de partículas. A las partículas que interfieren con signo *positivo* se les denomina *bosones*, y a las que interfieren con signo negativo, *fermiones*. Los bosones son el fotón, los mesones y el gravitón. Los fermiones son el electrón, el muon, los neutrinos, los nucleones y los bariones. Tenemos, entonces, que la amplitud para la dispersión de partículas idénticas son

Bosones:

$$(\text{Amplitud proceso directo}) + (\text{Amplitud proceso intercambiado}). \quad (4.1)$$

Fermiones:

$$(\text{Amplitud proceso directo}) - (\text{Amplitud proceso intercambiado}). \quad (4.2)$$

Para las partículas con espín —como los electrones— hay una complicación adicional. No sólo debemos especificar la ubicación de las partículas, sino también la dirección de sus espines. Es sólo para las partículas idénticas *con estados de espín idénticos*, que las amplitudes interfieren cuando las partículas se intercambian. Si piensan en la dispersión de haces no polarizados —los cuales son una mezcla de estados de espín diferentes— necesitan un poco más de aritmética.

Un problema interesante se presenta cuando hay dos o más partículas fuertemente ligadas. Por ejemplo, una partícula α contiene cuatro partículas —dos neutrones y dos protones—. Cuando dos partículas α se dispersan, hay varias posibilidades. Puede ser que durante la dispersión haya cierta amplitud para que uno de los neutrones pase de una partícula α a la otra, mientras que un neutrón de la otra partícula α haga lo contrario, de modo que las dos partículas α que salen de la dispersión no son las originales —ha habido un intercambio de un par de neutrones—. Véase la figura 4.2. La amplitud para la dispersión con

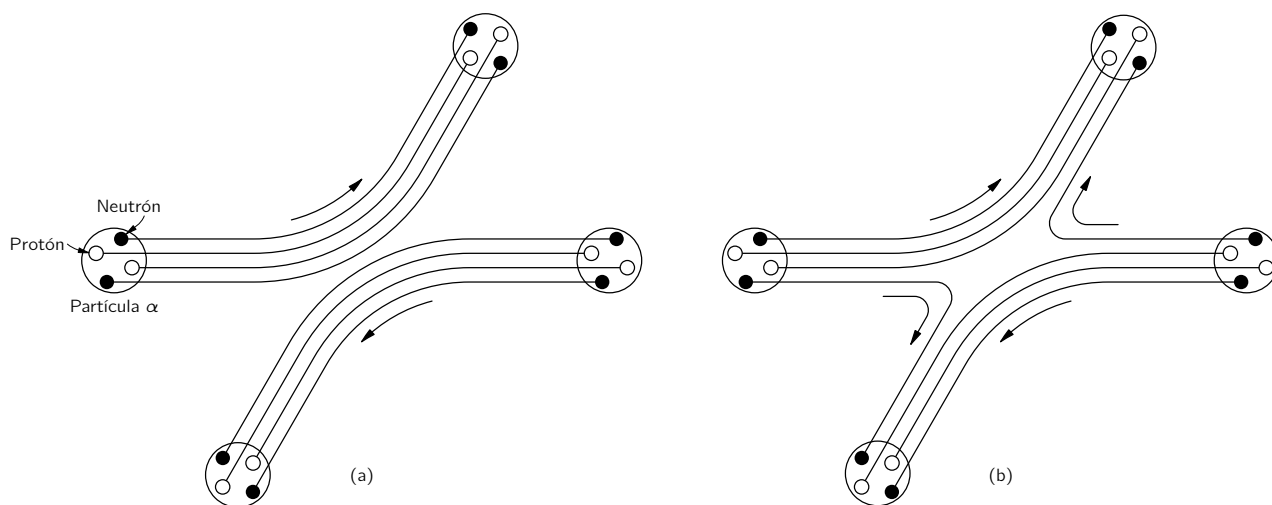


Figura 4-2. Dispersión de dos partículas α . En (a) las dos partículas mantienen su identidad; en (b) un neutrón se intercambia durante la colisión.

intercambio de un par de neutrones interferirá con la amplitud para la dispersión sin tal intercambio, y la interferencia debe ser con signo menos porque se han intercambiado un par de fermiones. Por otra parte, si la energía relativa de las dos partículas α es tan baja que éstas permanecen bastante alejadas una de otra —digamos, debido a la repulsión de Coulomb— y no hay ninguna probabilidad apreciable de que se intercambie alguna de las partículas internas, podemos considerar a la partícula α como un objeto simple, y no es necesario que nos preocupemos de sus detalles internos. En tal caso, sólo hay dos contribuciones a la amplitud para la dispersión. Ya sea que no hay intercambio, o los cuatro nucleones se intercambian en la dispersión. Debido a que los protones y los neutrones en la partícula α son fermiones, el intercambio de cualquier par invierte el signo de la amplitud para la dispersión. Mientras no haya cambios internos en las partículas α , intercambiar las dos partículas α es lo mismo que intercambiar cuatro pares de fermiones. Hay un cambio de signo para cada par, por lo que el resultado final es que las amplitudes se combinan con signo positivo. La partícula α se comporta como un bosón.

La regla es que los objetos compuestos, en los casos en que el objeto compuesto se puede considerar como un objeto simple, se comportan como fermiones o como bosones, dependiendo de si contienen un número impar o par de fermiones.

Todos los fermiones elementales que hemos mencionado —como el electrón, el protón, el neutrón y demás— tienen un espín $j = 1/2$. Si se juntan varios de esos fermiones para formar un objeto compuesto, el espín resultante puede ser entero o fraccionario. Por ejemplo, el isótopo común del helio, He^4 , el cual tiene dos neutrones y dos protones, tiene espín cero, mientras que el Li^7 , con tres protones y cuatro neutrones, tiene espín $3/2$. Después aprenderemos las reglas para componer el momentum angular; y ahora sólo mencionaremos que todo objeto compuesto que tiene *espín fraccionario* imita a un *fermión*, mientras que todo objeto compuesto con *espín entero* imita a un *bosón*.

Esto conduce a una pregunta interesante: ¿por qué las partículas con espín fraccionario son fermiones, cuyas amplitudes se suman con signo menos, mientras que las partículas con espín entero son bosones, cuyas amplitudes se suman con signo positivo? Pedimos disculpas por no poder darles una explicación sencilla. Pauli ha dado una explicación partiendo de argumentos complicados de teoría cuántica de campos y relatividad. Él ha demostrado que necesariamente las dos deben ir juntas, pero no hemos podido encontrar un modo de reproducir sus argumentos en un nivel sencillo. Parece ser uno de los pocos lugares de la física donde hay una regla que se puede enunciar de manera sencilla, pero para la cual nadie ha encontrado una explicación simple y sencilla. La explicación se encuentra en lo profundo de la mecánica cuántica relativista. Esto probablemente significa que no tenemos una comprensión completa del principio fundamental involucrado. Por el momento, sólo deberán aceptarla como una de las reglas del mundo.

4-2 Estados con dos bosones

En este momento nos gustaría discutir una consecuencia interesante de la regla de adición para bosones. Se refiere a su comportamiento cuando hay varias partículas presentes. Comenzamos considerando el caso en que dos bosones son dispersados desde dos dispersores diferentes. No nos preocuparemos por los detalles del mecanismo de dispersión. Sólo nos interesa lo que le ocurre a las partículas dispersadas. Supongamos que tenemos la situación que se muestra en la figura 4-3. La partícula a es dispersada hacia el estado 1. Por *estado* entendemos una dirección y una energía determinadas, o cualquier otra condición dada. La partícula b es dispersada hacia el estado 2. Queremos suponer que los estados 1 y 2 son casi iguales. (En realidad, lo que queremos averiguar es la amplitud para que las dos partículas sean dispersadas en direcciones idénticas, o estados idénticos; pero es mejor si primero pensamos en qué ocurre si los estados son casi iguales y luego vemos qué ocurre cuando se hacen idénticos.)

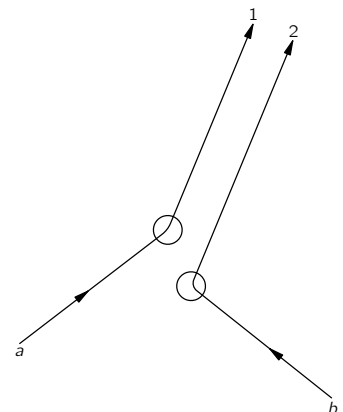


Figura 4-3. Dispersión doble con estados finales próximos.

Si sólo tenemos a la partícula a , entonces ésta tendría una cierta amplitud para ser dispersada en la dirección 1, digamos $\langle 1 | a \rangle$. Y la partícula b sola tendría la amplitud $\langle 2 | b \rangle$ para llegar a la dirección 2. Si las dos partículas no son idénticas, la amplitud para que las dos dispersiones ocurran al mismo tiempo es simplemente el producto

$$\langle 1 | a \rangle \langle 2 | b \rangle.$$

La probabilidad para tal evento es entonces

$$|\langle 1 | a \rangle \langle 2 | b \rangle|^2,$$

que es igual a

$$|\langle 1 | a \rangle|^2 |\langle 2 | b \rangle|^2.$$

Para ahorrar escritura en esta discusión, a veces diremos

$$\langle 1 | a \rangle = a_1, \quad \langle 2 | b \rangle = b_2.$$

Entonces la probabilidad para la dispersión doble es

$$|a_1|^2 |b_2|^2.$$

También podría ocurrir que la partícula b sea dispersada en la dirección 1, mientras que la partícula a fuera en la dirección 2. La amplitud para que este proceso ocurra es

$$\langle 2 | a \rangle \langle 1 | b \rangle,$$

y la probabilidad para este evento es

$$|\langle 2 | a \rangle \langle 1 | b \rangle|^2 = |a_2|^2 |b_1|^2.$$

Imaginemos ahora que tenemos un par de contadores pequeñísimos que reciben a las dos partículas dispersadas. La probabilidad P_2 para que reciban a las dos partículas juntas es sólo la suma

$$P_2 = |a_1|^2 |b_2|^2 + |a_2|^2 |b_1|^2. \quad (4.3)$$

Ahora supongamos que las direcciones 1 y 2 son muy cercanas una de la otra. Esperamos que a varíe de manera continua con la dirección, de modo que a_1 y a_2 se aproximen a medida que 1 y 2 se acercan. Si están lo suficientemente cerca, las amplitudes a_1 y a_2 serán iguales. Podemos decir que $a_1 = a_2$ y llamar a las dos simplemente a ; análogamente, decimos que $b_1 = b_2 = b$. Entonces tenemos que

$$P_2 = 2|a|^2 |b|^2. \quad (4.4)$$

Sin embargo, supongamos ahora que a y b son bosones idénticos. Entonces, el proceso en que a va hacia 1 y b va hacia 2 no se puede distinguir del proceso intercambiado, en que a va hacia 2 y b va hacia 1. En este caso las *amplitudes* para los dos procesos diferentes pueden interferir. La amplitud *total* para obtener una partícula en cada contador es

$$\langle 1 | a \rangle \langle 2 | b \rangle + \langle 2 | a \rangle \langle 1 | b \rangle. \quad (4.5)$$

Y la probabilidad para que obtengamos un par es el cuadrado de la norma de esta amplitud,

$$P_2 = |a_1 b_2 + a_2 b_1|^2 = 4|a|^2 |b|^2. \quad (4.6)$$

Obtenemos el resultado de que es *dos veces más probable* encontrar dos bosones idénticos dispersados hacia el mismo estado, *de lo que se obtiene suponiendo que las partículas son diferentes*.

Aunque hemos considerado que se observa a las dos partículas en contadores separados, esto no es necesario —pues podemos verlo de la manera siguiente—. Imaginemos que ambas direcciones 1 y 2 llevan a las partículas a un *solo* contador pequeño que está a cierta distancia. Definimos a la dirección 1 diciendo que

apunta hacia el elemento de superficie dS_1 del contador. La dirección 2 apunta hacia el elemento de superficie dS_2 del contador. (Imaginemos que el contador presenta una superficie perpendicular a la línea de las dispersiones.) Ahora no podemos dar una probabilidad para que una partícula vaya en una dirección precisa, o hacia un punto *específico* del espacio. Eso es imposible —la probabilidad para cualquier dirección exacta es cero—. Cuando queremos ser tan específicos, tenemos que definir nuestras amplitudes de modo tal, que den la probabilidad para llegar a una *unidad* de superficie de un contador. Supongamos que sólo tenemos a la partícula a ; ésta tendría una cierta amplitud de dispersión en la dirección 1. Definamos a $\langle 1 | a \rangle = a_1$ como la amplitud para que a se disperse hacia una *unidad* de superficie del contador en la dirección 1. En otras palabras, se elige la escala de a_1 —decimos que está “normalizada”—, de modo que la probabilidad para que se disperse hacia un elemento de superficie dS_1 es

$$|\langle 1 | a \rangle|^2 dS_1 = |a_1|^2 dS_1. \quad (4.7)$$

Si la superficie total de nuestro contador es ΔS , y hacemos que dS_1 recorra esta superficie, la probabilidad total para que la partícula a se disperse hacia el contador es

$$\int_{\Delta S} |a_1|^2 dS_1. \quad (4.8)$$

Como hicimos antes, suponemos que el contador es suficientemente pequeño, como para que la amplitud a_1 no varíe de manera significativa sobre la superficie del contador; a_1 es entonces una amplitud constante que podemos llamar a . Entonces la probabilidad para que la partícula a se disperse hacia algún punto del contador es

$$p_a = |a|^2 \Delta S. \quad (4.9)$$

Del mismo modo, tenemos la probabilidad para que la partícula b —cuando está sola— se disperse hacia algún elemento de superficie, digamos dS_2 , es

$$|b_2|^2 dS_2.$$

(Usamos dS_2 en vez de dS_1 , porque más tarde necesitaremos que a y b vayan en diferentes direcciones.) De nuevo, hacemos que b_2 sea igual a la amplitud constante b , entonces la probabilidad para que la partícula b sea contada en el detector es

$$p_b = |b|^2 \Delta S. \quad (4.10)$$

Ahora, cuando ambas partículas están presentes, la probabilidad para que a se disperse hacia dS_1 y b hacia dS_2 es

$$|a_1 b_2|^2 dS_1 dS_2 = |a|^2 |b|^2 dS_1 dS_2. \quad (4.11)$$

Si queremos la probabilidad para que *ambas* a y b entren en el contador, integramos ambas dS_1 y dS_2 sobre ΔS y encontramos que

$$P_2 = |a|^2 |b|^2 (\Delta S)^2. \quad (4.12)$$

Notemos que esto es justamente igual a $p_a \cdot p_b$, tal como podría esperarse, suponiendo que las partículas a y b actúan independientemente una de la otra.

No obstante, cuando las dos partículas son idénticas, hay dos posibilidades que no se pueden distinguir para cada par de elementos de superficie dS_1 y dS_2 . Que la partícula a entre en dS_2 y la partícula b en dS_1 es indistinguible de que la partícula a entre en dS_1 y la b en dS_2 , por lo que las amplitudes para estos dos procesos van a interferir. (Cuando teníamos dos partículas *diferentes* —aunque *en realidad* no nos importaba cuál partícula iba a tal lugar del contador— podíamos, *en principio*, distinguirlas; por lo que no había interferencia. En el caso de partículas idénticas, no podemos distinguirlas, ni aún en principio.) Debemos escribir, entonces, que la probabilidad para que las dos partículas lleguen a dS_1 y dS_2 es

$$|a_1 b_2 + a_2 b_1|^2 dS_1 dS_2. \quad (4.13)$$

Ahora, sin embargo, al integrar sobre la superficie del contador, debemos tener cuidado. Si hacemos que dS_1 y dS_2 recorran toda la superficie ΔS , contaríamos *dos veces* cada parte de la superficie, ya que (4.13) contiene todo lo que puede ocurrir con cualquier par de elementos de superficie dS_1 y dS_2 .[†] Aún podemos integrar de esa manera, si compensamos el doble conteo dividiendo el resultado entre 2. Obtenemos entonces que, para bosones idénticos, P_2 es

$$P_2(\text{bosón}) = \frac{1}{2} \{4|a|^2|b|^2 (\Delta S)^2\} = 2|a|^2|b|^2 (\Delta S)^2. \quad (4.14)$$

Una vez más, esto es el doble de lo que obtuvimos en la ecuación (4.12) para partículas distinguibles.

Si imaginamos por un momento, que sabíamos que el canal b ya había mandado su partícula en una dirección determinada, podemos decir que la *probabilidad* para que una segunda partícula vaya en la misma dirección, es dos veces mayor de lo que habríamos esperado de haberla calculado como un evento independiente. Es propiedad de los bosones que, si ya hay una partícula en alguna condición, la *probabilidad* para tener una segunda en la misma condición es dos veces mayor de lo que sería si la primera partícula no estuviera ya ahí. Este hecho a menudo se enuncia de la manera siguiente: si ya hay un bosón en un estado determinado, la amplitud correspondiente a poner otro idéntico sobre él es $\sqrt{2}$ veces mayor que si éste no estuviera ahí. (Ésta no es la manera apropiada de enunciar el resultado desde el punto de vista físico que hemos adoptado, pero si se usa sistemáticamente como regla, dará el resultado correcto.)

4-3 Estados con n bosones

Extendamos nuestro resultado a un caso en el cual hay n partículas presentes. Imaginemos la situación mostrada en la figura 4-4. Tenemos n partículas $a, b, c, \dots$, que se dispersan y terminan en las direcciones $1, 2, 3, \dots, n$. Todas las n direcciones apuntan hacia un contador pequeño muy lejano. De la misma forma que en la sección anterior, elegimos normalizar todas las amplitudes, de modo que la probabilidad para que cada partícula actuando sola vaya a un elemento de superficie dS del contador es

$$|\langle \rangle|^2 dS.$$

Primero, supongamos que todas las partículas son distinguibles; entonces la probabilidad para que se cuenten n partículas juntas en n elementos de superficie diferentes es

$$|a_1 b_2 c_3 \dots|^2 dS_1 dS_2 dS_3 \dots \quad (4.15)$$

De nuevo, consideramos que las amplitudes no dependen de dónde está ubicado dS en el contador (que supusimos pequeño), y las llamamos de manera sencilla $a, b, c, \dots$. La probabilidad (4.15) se convierte en

$$|a|^2 |b|^2 |c|^2 \dots dS_1 dS_2 dS_3 \dots \quad (4.16)$$

Cuando integramos cada dS sobre la superficie ΔS del contador, tenemos que P_n (diferentes), la probabilidad para contar n partículas distintas a la vez, es

$$P_n(\text{diferentes}) = |a|^2 |b|^2 |c|^2 \dots (\Delta S)^n. \quad (4.17)$$

Éste es sólo el producto de las probabilidades para que cada partícula entre al contador por separado. Todas actúan de manera independiente —la probabilidad para que entre una no depende de cuántas más están entrando—.

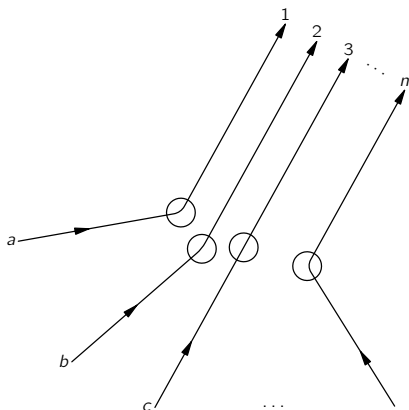


Figura 4-4. La dispersión de n partículas con estados finales próximos.

[†] En (4.11) intercambiar dS_1 y dS_2 produce un evento diferente, por lo que los dos elementos de superficie deben recorrer la superficie total del contador. En (4.13) tratamos a dS_1 y dS_2 como un *par* e incluimos en él todo lo que pueda pasar. Si las integrales incluyen lo que ocurre cuando se intercambian dS_1 y dS_2 , todo se cuenta dos veces.

Ahora supongamos que todas las partículas son bosones idénticos. Para cada conjunto de direcciones $1, 2, 3, \dots$ hay muchas posibilidades indistinguibles. Si sólo hubiera, por ejemplo, tres partículas, tendríamos las siguientes posibilidades:

$$\begin{array}{lll} a \rightarrow 1 & a \rightarrow 1 & a \rightarrow 2 \\ b \rightarrow 2 & b \rightarrow 3 & b \rightarrow 1 \\ c \rightarrow 3 & c \rightarrow 2 & c \rightarrow 3 \\ \\ a \rightarrow 2 & a \rightarrow 3 & a \rightarrow 3 \\ b \rightarrow 3 & b \rightarrow 1 & b \rightarrow 2 \\ c \rightarrow 1 & c \rightarrow 2 & c \rightarrow 1 \end{array}$$

Hay seis combinaciones diferentes. Con n partículas hay $n!$ diferentes, aunque *indistinguibles*, posibilidades para las cuales debemos sumar las amplitudes. La probabilidad para que se cuente n partículas en n elementos de superficie es entonces

$$|a_1 b_2 c_3 \cdots + a_1 b_3 c_2 \cdots + a_2 b_1 c_3 \cdots + a_2 b_3 c_1 \cdots + \text{etc.} + \text{etc.}|^2 dS_1 dS_2 dS_3 \cdots dS_n. \quad (4.18)$$

Una vez más, supongamos que todas las direcciones están muy cerca, y que podemos decir que $a_1 = a_2 = \cdots = a_n = a$, y análogamente para $b, c, \dots$; la probabilidad (4.18) se convierte en

$$|n! abc \cdots|^2 dS_1 dS_2 \cdots dS_n. \quad (4.19)$$

Cuando integramos cada dS sobre la superficie ΔS del contador, cada producto posible de elementos de superficie es contado $n!$ veces; compensamos esto dividiendo entre $n!$ y obtenemos

$$P_n(\text{bosón}) = \frac{1}{n!} |n! abc \cdots|^2 (\Delta S)^n$$

o

$$P_n(\text{bosón}) = n! |n! abc \cdots|^2 (\Delta S)^n. \quad (4.20)$$

Al comparar este resultado con la ecuación (4.17), vemos que la probabilidad para contar n bosones a la vez es $n!$ veces mayor de lo que obtendríamos suponiendo que las partículas fueran distinguibles. Podemos resumir nuestro resultado así:

$$P_n(\text{bosón}) = n! P_n(\text{diferentes}). \quad (4.21)$$

Por lo tanto, la probabilidad, en el caso de bosones, es mayor por $n!$ de lo que se obtendría suponiendo que las partículas actúan de manera independiente.

Podemos entender mejor el significado de esto si hacemos la siguiente pregunta: ¿cuál es la probabilidad para que un bosón vaya a un estado determinado cuando *ya hay otros* n presentes? Llamemos a la partícula recién agregada w . Si tenemos $(n+1)$ partículas, incluyendo a w , la ecuación (4.20) se convierte en

$$P_{n+1}(\text{bosón}) = (n+1)! |abc \cdots w|^2 (\Delta S)^{n+1}. \quad (4.22)$$

La podemos escribir en la forma

$$P_{n+1}(\text{bosón}) = \{(n+1)|w|^2 \Delta S\} n! |abc \cdots|^2 (\Delta S)^n$$

o

$$P_{n+1}(\text{bosón}) = (n+1)|w|^2 \Delta S P_n(\text{bosón}). \quad (4.23)$$

Podemos interpretar este resultado de la siguiente manera: el número $|w|^2 \Delta S$ es la probabilidad para que la partícula w entre al detector si no hay otras partículas presentes; $P_n(\text{bosón})$ es la probabilidad para que haya otros n bosones presentes. Entonces la ecuación (4.23) dice que, *cuando* están presentes otros n bosones idénticos, la probabilidad para que *una* partícula *más* llegue al mismo estado *aumenta* por un factor de $(n+1)$. La probabilidad para obtener un bosón, donde ya hay n , es $(n+1)$ veces mayor de lo que sería si no hubiera ninguno antes. La *presencia* de las otras partículas aumenta la probabilidad para obtener una más.

4-4 Emisión y absorción de fotones

En nuestra discusión, hemos hablado de un proceso como la dispersión de partículas α . Pero eso no es necesario; podríamos haber estado hablando de la creación de partículas, como en la emisión de luz, por ejemplo. Cuando se emite luz, se “crea” un fotón. En este caso, no necesitamos las líneas entrantes de la figura 4-4; simplemente, podemos considerar que hay algunos átomos emitiendo n fotones, como en la figura 4-5. Por lo tanto, nuestro resultado también se puede enunciar así: *la probabilidad para que un átomo emita un fotón en un estado final determinado, aumenta por un factor $(n + 1)$ si ya hay n fotones en el mismo estado.*

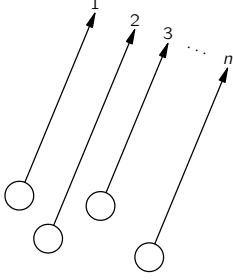


Figura 4-5. Creación de n fotones en estados próximos.

A la gente le gusta resumir este resultado, diciendo que la *amplitud* para la emisión de un fotón aumenta por un factor $\sqrt{n + 1}$ cuando ya hay n fotones. Ésta es, por supuesto, otra manera de decir lo mismo, si se sobreentiende que sólo hay que elevar al cuadrado la norma de esa amplitud para obtener la probabilidad.

Por lo general, en mecánica cuántica, es cierto que la amplitud para ir de cualquier estado ϕ a cualquier otro estado χ es el complejo conjugado de la amplitud para ir de χ a ϕ :

$$\langle \chi | \phi \rangle = \langle \phi | \chi \rangle^*. \quad (4.24)$$

Aprenderemos sobre esta ley un poco más adelante; por el momento supondremos simplemente que es verdadera. Podemos usarla para encontrar cómo los fotones son dispersados o absorbidos *fuera* de un estado determinado. Sabemos que la amplitud para que un fotón sea agregado a algún estado, digamos i , cuando ya hay n fotones presentes es, digamos,

$$\langle n + 1 | n \rangle = \sqrt{n + 1} a, \quad (4.25)$$

donde $a = \langle i | a \rangle$ es la amplitud cuando no hay otros fotones presentes. Usando la ecuación (4.24), la amplitud para ir a la inversa —de $(n + 1)$ fotones a n — es

$$\langle n | n + 1 \rangle = \sqrt{n + 1} a^*. \quad (4.26)$$

Ésta no es la forma en que la gente lo expresa comúnmente; no les gusta pensar en ir de $(n + 1)$ a n , sino que prefieren siempre comenzar con n fotones presentes. Dicen que la amplitud para absorber un fotón cuando hay n presentes —en otras palabras, ir de n a $(n - 1)$ — es

$$\langle n - 1 | n \rangle = \sqrt{n} a^*. \quad (4.27)$$

que es, por supuesto, igual a la ecuación (4.26). Luego tienen problemas tratando de recordar si usar $\sqrt{n}$ o $\sqrt{n + 1}$. La manera de recordar es la siguiente: el factor es siempre la raíz cuadrada del número mayor de fotones presentes, ya sea antes o después de la reacción. Las ecuaciones (4.25) y (4.26) demuestran que la ley es realmente simétrica —sólo no aparece simétrica cuando se escribe como en la ecuación (4.27)—.

Hay muchas consecuencias físicas de estas nuevas reglas; queremos describir una de ellas que tiene que ver con la emisión de la luz. Imaginemos una situación en la cual los fotones están contenidos dentro de una caja —pueden imaginar una caja cuyas paredes son espejos—. Digamos que en la caja tenemos n fotones, todos en el mismo estado —la misma frecuencia, dirección y polarización— así que no se les puede distinguir, y que también hay un átomo en la caja que puede emitir otro fotón en el mismo estado. Entonces la probabilidad para que emita un fotón es

$$(n + 1)|a|^2, \quad (4.28)$$

y la probabilidad para que absorba un fotón es

$$n|a|^2, \quad (4.29)$$

donde $|a|^2$ es la probabilidad que tendría para emitir si no hubiera fotones presentes. Ya hemos discutido esta regla de un modo algo diferente en el capítulo

42 del volumen I. La ecuación (4.29) dice que la probabilidad para que un átomo *absorba* un fotón, y haga una transición a un estado de energía mayor, es proporcional a la intensidad de la luz que brilla sobre él. Pero, como primero Einstein señaló, la rapidez con la que un átomo hace una transición *hacia abajo* tiene dos partes. Está la probabilidad para que haga una transición espontánea $|a|^2$, más la probabilidad para una transición inducida $n|a|^2$, que es proporcional a la intensidad de la luz —o sea, al número de fotones presentes—. Además, como Einstein dijo, los coeficientes de absorción y de emisión inducida son iguales, y están relacionados con la probabilidad para la emisión espontánea. De aquí aprendemos que, si la intensidad de la luz se mide en términos del número de fotones presentes (en lugar de la energía por unidad de área, y por segundo), los coeficientes de absorción, de emisión inducida y de emisión espontánea son iguales. Éste es el significado de la relación entre los coeficientes de Einstein A y B que aparecen en el capítulo 42, volumen I, ecuación (42.18).

4-5 El espectro de cuerpo negro

Nos gustaría usar nuestras reglas para bosones para discutir de nuevo el espectro de la radiación de cuerpo negro (véase capítulo 42, volumen I). Lo haremos calculando cuántos fotones hay en una caja si la radiación está en equilibrio térmico con algunos átomos de la caja. Supongamos que, para cada frecuencia de luz ω , hay un cierto número N de átomos, los cuales tienen dos estados de energía separados por $\Delta E = \hbar\omega$. Véase figura 4-6. Llamaremos al estado de menor energía el estado “base”, y al de mayor energía el estado “excitado”. Sean N_g y N_e los números promedio de átomos en los estados base y excitado; entonces, en equilibrio térmico a la temperatura T , sabemos de mecánica estadística que

$$\frac{N_e}{N_g} = e^{-\Delta E/kT} = e^{-\hbar\omega/kT}. \quad (4.30)$$

Cada átomo en el estado base puede absorber un fotón y pasar al estado excitado, y cada átomo en el estado excitado puede emitir un fotón y pasar al estado base. En el equilibrio, las rapidezces de ambos procesos deben ser las mismas. Las rapidezces son proporcionales a la probabilidad para el evento y al número de átomos presentes. Llamemos $\bar{n}$ al número promedio de fotones presentes en un estado determinado con frecuencia ω . Entonces la rapidez de absorción de ese estado es $N_g\bar{n}|a|^2$, y el grado de emisión hacia ese estado es $N_e(\bar{n} + 1)|a|^2$. Igualando las dos rapidezces obtenemos:

$$N_g\bar{n} = N_e(\bar{n} + 1). \quad (4.31)$$

Al combinar este resultado con la ecuación (4.30), obtenemos

$$\frac{\bar{n}}{\bar{n} + 1} = e^{-\hbar\omega/kT}.$$

Al resolver para $\bar{n}$, obtenemos

$$\bar{n} = \frac{1}{e^{\hbar\omega/kT} - 1}, \quad (4.32)$$

que es el número promedio de fotones en cualquier estado de frecuencia ω , para una cavidad en equilibrio térmico. Como cada fotón tiene una energía $\hbar\omega$, la energía de los fotones en un estado determinado es $\bar{n}\hbar\omega$, o

$$\frac{\hbar\omega}{e^{\hbar\omega/kT} - 1}. \quad (4.33)$$

Como observación, ya habíamos encontrado una ecuación similar en otro contexto [capítulo 41, volumen I, ecuación (41.15)]. Ustedes recuerdan que para cualquier oscilador armónico —tal como una pesa suspendida de un resorte— los

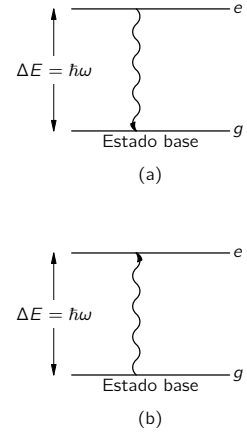


Figura 4-6. Radiación y absorción de un fotón con frecuencia ω .

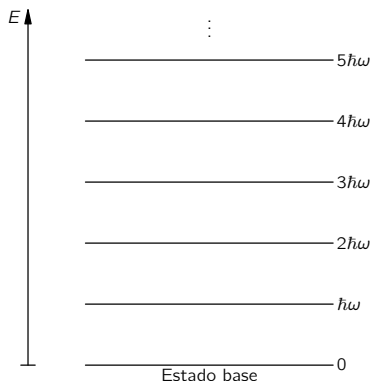


Figura 4-7. Niveles de energía para un oscilador armónico.

niveles de energía cuánticos están igualmente distribuidos con una separación $\hbar\omega$, como se ha dibujado en la figura 4-7. Si llamamos a la energía del n -ésimo nivel $n\hbar\omega$, encontramos que la energía media de ese oscilador también está dada por la ecuación (4.33). Sin embargo, esta ecuación se obtuvo aquí para fotones, contando partículas, y da los mismos resultados. Esto es uno de los maravillosos milagros de la mecánica cuántica. Si empezamos considerando un tipo de estado o condición para bosones que no interactúan entre sí (hemos supuesto que los fotones no interactúan entre sí), y luego consideramos que en este estado se pueden poner ya sean cero, o una, o dos, ... hasta cualquier número n de partículas, encontramos que este sistema, para todos los fines de la mecánica cuántica, se comporta exactamente como un oscilador armónico. Por este tipo de oscilador, entendemos un sistema dinámico, como una pesa suspendida de un resorte o una onda estacionaria en una cavidad resonante. Y ésa es la razón por la cual es posible representar el campo electromagnético por medio de partículas fotones. Desde un punto de vista, podemos analizar el campo electromagnético dentro de una caja o cavidad en términos de muchos osciladores armónicos, tratando cada modo de oscilación, de acuerdo con la mecánica cuántica, como un oscilador armónico. Desde otro punto de vista, podemos analizar la misma situación física en términos de bosones idénticos. Y los resultados de ambas maneras de trabajar *siempre concuerdan exactamente*. No hay manera de decidir si el campo electromagnético en realidad se debe describir como un oscilador armónico cuantizado, o dando la cantidad de fotones que hay en cada condición. Los dos puntos de vista resultan ser matemáticamente idénticos. En adelante, podemos hablar ya sea del número de fotones en un estado determinado dentro de una caja, o bien del número de niveles de energía asociado con un modo particular de oscilación del campo electromagnético. Son dos maneras de decir lo mismo. También es válido el mismo argumento para fotones en el espacio libre. Son equivalentes a las oscilaciones de una cavidad cuyas paredes se han encojido hasta el infinito.

Hemos calculado la energía promedio en cualquier modo particular dentro de una caja a la temperatura T ; sólo necesitamos una cosa más para obtener la ley de radiación de cuerpo negro: necesitamos saber cuántos modos hay para cada energía. (Suponemos que para cada modo hay algunos átomos dentro de la caja —o en las paredes— con niveles de energía que pueden radiar en ese modo, por lo que cada uno puede llegar al equilibrio térmico.) La ley de radiación de cuerpo negro por lo general se enuncia dando la energía por unidad de volumen que transporta la luz en un pequeño intervalo de frecuencia entre ω y $\omega + \Delta\omega$. Así que necesitamos saber cuántos modos hay en una caja con frecuencias dentro del intervalo $\Delta\omega$. Aunque esta pregunta se presenta continuamente en la mecánica cuántica, es simplemente una pregunta clásica sobre ondas estacionarias.

Obtendremos la respuesta sólo para una caja rectangular. Ocurre lo mismo para una caja de cualquier forma, pero el cálculo es muy complicado para el caso arbitrario. Además, sólo estamos interesados en una caja cuyas dimensiones son muy grandes, comparadas con la longitud de onda de la luz. Entonces hay miles de millones y miles de millones de modos; habrá muchos en cualquier intervalo pequeño de frecuencias $\Delta\omega$, por lo que podemos hablar del “número promedio” en cualquier $\Delta\omega$ a la frecuencia ω . Empecemos por preguntar cuántos modos hay en un caso unidimensional —como las ondas en una cuerda tensa—. Ustedes saben que cada modo es una onda sinusoidal que tiene que anularse en ambos extremos; en otras palabras, tiene que haber un múltiplo entero de la mitad de las longitudes de onda en la longitud de la línea, como se muestra en la figura 4-8. Preferimos usar el número de onda $k = 2\pi/\lambda$; llamando k_j al número de onda del j -ésimo modo, tenemos que

$$k_j = \frac{j\pi}{L}, \quad (4.34)$$

donde j es cualquier número entero. La separación δk entre modos sucesivos es

$$\delta k = k_{j+1} - k_j = \frac{\pi}{L}.$$

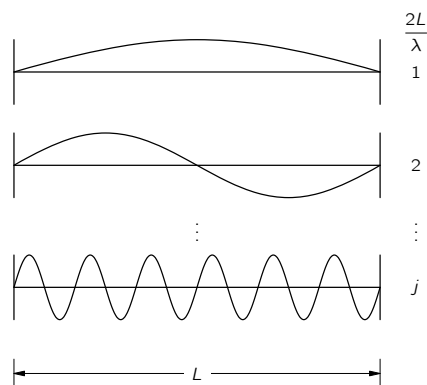


Figura 4-8. Modos de ondas estacionarias en una dimensión.

Queremos suponer que kL es tan grande que, en un pequeño intervalo Δk , hay muchos modos. Llamando $\Delta \mathfrak{N}$ al número de modos que hay en el intervalo Δk , tenemos:

$$\Delta \mathfrak{N} = \frac{\Delta k}{\delta k} = \frac{L}{\pi} \Delta k. \quad (4.35)$$

Ahora bien, los físicos teóricos, que trabajan en mecánica cuántica, por lo general prefieren decir que hay la mitad de modos; ellos escriben:

$$\Delta \mathfrak{N} = \frac{L}{2\pi} \Delta k. \quad (4.36)$$

Queremos explicar por qué. Por lo general, a ellos les gusta pensar en términos de ondas viajeras —unas que van hacia la derecha (con k positivo) y otras que van hacia la izquierda (con k negativo)—. Pero un “modo” es una onda *estacionaria*, que es la suma de dos ondas, una hacia cada dirección. En otras palabras, consideran que cada onda estacionaria contiene dos “estados” fotónicos distintos. Por ello, si preferimos que $\Delta \mathfrak{N}$ signifique el número de estados fotónicos para un k determinado (donde k ahora toma valores positivos y negativos), debemos tomar la mitad de $\Delta \mathfrak{N}$. (Todas las integrales deben ir ahora de $k = -\infty$ a $k = +\infty$, y el número total de estados hasta cualquier valor absoluto de k saldrá correctamente.) Por supuesto, no estamos describiendo muy bien las ondas estacionarias, pero estamos contando modos en una forma consistente.

Ahora queremos extender este resultado a tres dimensiones. Una onda estacionaria dentro de una caja rectangular debe tener un múltiplo entero de medias ondas *para cada eje*. La situación para dos dimensiones se muestra en la figura 4-9. Cada dirección y frecuencia es descrita por un vector de onda $\mathbf{k}$, cuyas componentes x , y y z deben satisfacer ecuaciones como la ecuación (4.34). Entonces tenemos

$$\begin{aligned} k_x &= \frac{j_x \pi}{L_x}, \\ k_y &= \frac{j_y \pi}{L_y}, \\ k_z &= \frac{j_z \pi}{L_z}. \end{aligned}$$

El número de modos con k_x dentro de intervalo Δk_x es, como antes,

$$\frac{L_x}{2\pi} \Delta k_x,$$

y análogamente para Δk_y y Δk_z . Si llamamos $\Delta \mathfrak{N}(\mathbf{k})$ al número de modos para un vector de onda $\mathbf{k}$, cuya componente x está entre k_x y $k_x + \Delta k_x$, cuya componente y está entre k_y y $k_y + \Delta k_y$, y cuya componente z está entre k_z y $k_z + \Delta k_z$, entonces

$$\Delta \mathfrak{N}(\mathbf{k}) = \frac{L_x L_y L_z}{(2\pi)^3} \Delta k_x \Delta k_y \Delta k_z. \quad (4.37)$$

El producto $L_x L_y L_z$ es igual al volumen V de la caja. Así que llegamos al resultado importante de que, para frecuencias altas (longitudes de onda pequeñas comparadas con las dimensiones), el número de modos en una cavidad es proporcional al volumen V de la caja y al “volumen en el k -espacio” $\Delta k_x \Delta k_y \Delta k_z$. Este resultado aparece repetidamente en muchos problemas y se debe memorizar:

$$d\mathfrak{N}(\mathbf{k}) = V \frac{d^3 \mathbf{k}}{(2\pi)^3}. \quad (4.38)$$

Aunque no lo hayamos probado, este resultado es independiente de la forma de la caja.

Aplicaremos este resultado para encontrar el número de modos fotónicos para fotones con frecuencias en el intervalo $\Delta \omega$. Sólo estamos interesados en la energía

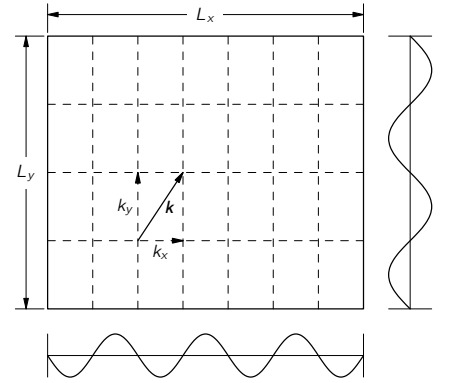


Figura 4-9. Modos de ondas estacionarias en dos dimensiones.

asociada con varios modos —mas no en la dirección de las ondas—. Nos gustaría saber el número de modos en un intervalo determinado de frecuencias. En el vacío la magnitud de $\mathbf{k}$ está relacionada con la frecuencia por

$$|\mathbf{k}| = \frac{\omega}{c}. \quad (4.39)$$

Por lo tanto, en un intervalo de frecuencias $\Delta\omega$, éstos son todos los modos que corresponden a los $\mathbf{k}$ con *magnitud* entre k y $k + \Delta k$, independiente de la dirección. El “volumen en el k -espacio” entre k y $k + \Delta k$ es una cáscara esférica de volumen

$$4\pi k^2 \Delta k.$$

Entonces el número de modos es

$$\Delta\mathfrak{N}(\omega) = \frac{V 4\pi k^2 \Delta k}{(2\pi)^3}. \quad (4.40)$$

Sin embargo, como estamos interesados en frecuencias, sustituimos $k = \omega/c$ para obtener

$$\Delta\mathfrak{N}(\omega) = \frac{V 4\pi \omega^2 \Delta\omega}{(2\pi)^3 c^3}. \quad (4.41)$$

Hay una complicación adicional. Si estamos hablando de los modos de una onda electromagnética, cada vector de onda $\mathbf{k}$ puede estar en cualquiera de dos polarizaciones (perpendiculares entre sí). Como estos modos son independientes, debemos —para la luz— duplicar el número de modos. Entonces tenemos:

$$\Delta\mathfrak{N}(\omega) = \frac{V \omega^2 \Delta\omega}{\pi^2 c^3} \quad (\text{para la luz}). \quad (4.42)$$

Hemos demostrado, ecuación (4.33), que cada modo (o cada “estado”) tiene en promedio la energía

$$\bar{n}\hbar\omega = \frac{\hbar\omega}{e^{\hbar\omega/kT} - 1}.$$

Multiplicando esta ecuación por el número de modos, obtenemos la energía ΔE en los modos que están en el intervalo $\Delta\omega$:

$$\Delta E = \frac{\hbar\omega}{e^{\hbar\omega/kT} - 1} \frac{V \omega^2 \Delta\omega}{\pi^2 c^3}. \quad (4.43)$$

Ésta es la ley para el espectro de frecuencias de la radiación de cuerpo negro, que ya encontramos en el capítulo 41 del volumen I. El espectro está graficado en la figura 4-10. Ahora ven que la respuesta depende del hecho de que los fotones son bosones, los cuales tienen una tendencia de ir todos al mismo estado (porque la amplitud para hacerlo es grande). Como recordarán, fue el estudio de Planck sobre el espectro de cuerpo negro (que era un misterio para la física clásica), y su descubrimiento de la fórmula (4.43) lo que originó toda la mecánica cuántica.

4-6 El helio líquido

A temperaturas bajas, el helio líquido tiene muchas propiedades raras que, lamentablemente, no tenemos tiempo de describir en detalle en este momento, pero muchas de ellas provienen del hecho de que el átomo de helio es un bosón. Una de las cosas es que el helio líquido fluye sin resistencia viscosa alguna. Es, en realidad, el agua “seca” ideal de la que hemos estado hablando en capítulos anteriores —siempre y cuando las velocidades sean suficientemente bajas—. La razón es la siguiente. Para que un líquido tenga viscosidad, debe haber pérdidas internas de energía; debe haber alguna manera de que una parte del líquido tenga un movimiento diferente al del resto del líquido. Esto significa que debe ser posible empujar algunos átomos hacia estados diferentes de los estados ocupados por otros átomos. Pero a temperaturas suficientemente bajas, cuando los movimientos térmicos se hacen muy pequeños, todos los átomos tratan de ir

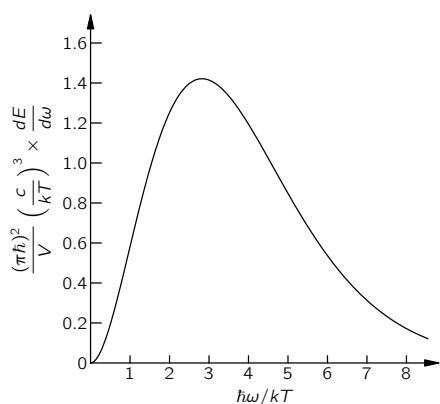


Figura 4-10. El espectro de frecuencias de la radiación en una cavidad en equilibrio térmico, el espectro de “cuerpo negro”.

al mismo estado. Por ello, si algunos se están moviendo juntos, entonces todos los átomos tratan de moverse juntos hacia el mismo estado. Hay una especie de rigidez en el movimiento, y es difícil dividir el movimiento en patrones irregulares de turbulencia, como ocurriría, por ejemplo, con partículas independientes. Así que en un líquido de bosones, hay una tendencia fuerte en todos los átomos de ir al mismo estado —que está representada por el factor $\sqrt{n+1}$ que encontramos antes—. (Para una botella de helio líquido n es, claramente, un número ¡muy grande!) Este movimiento cooperativo no ocurre a temperaturas altas, porque entonces hay suficiente energía térmica como para poner a los diversos átomos en varios estados más altos diferentes. Pero, a una temperatura suficientemente baja, repentinamente llega un momento en que todos los átomos de helio tratan de ir al mismo estado. El helio se convierte en un superfluido. Además, este fenómeno sólo aparece en el isótopo del helio que tiene peso atómico 4. En el caso de isótopo de helio con peso atómico 3, los átomos individuales son fermiones, y el líquido es un fluido normal. Como la superfluidez ocurre sólo con He^4 , es evidentemente un efecto cuántico —debido a la naturaleza bosónica de la partícula α —.

4-7 El principio de exclusión

Los fermiones actúan de un modo completamente diferente. Veamos lo que ocurre si tratamos de poner dos fermiones en el mismo estado. Volveremos a nuestro ejemplo original y preguntemos por la amplitud para que dos fermiones idénticos sean dispersados en casi exactamente la misma dirección. La amplitud para que la partícula a vaya en la dirección 1, y la partícula b en la dirección 2 es

$$\langle 1 | a \rangle \langle 2 | b \rangle,$$

mientras que la amplitud para cuando las direcciones se intercambian es

$$\langle 2 | a \rangle \langle 1 | b \rangle.$$

Como tenemos fermiones, la amplitud para el proceso es la diferencia de estas dos amplitudes:

$$\langle 1 | a \rangle \langle 2 | b \rangle - \langle 2 | a \rangle \langle 1 | b \rangle. \quad (4.44)$$

Digamos que por “dirección 1” entendemos que la partícula no sólo tiene una cierta dirección, sino también una dirección determinada para su espín, y que la “dirección 2” es casi igual a la dirección 1, y corresponde a la *misma* dirección para el espín. Entonces $\langle 1 | a \rangle$ y $\langle 2 | a \rangle$ son casi iguales. (Esto no sería necesariamente cierto si los estados 1 y 2 no tuvieran el mismo espín, porque podría haber alguna razón para que la amplitud dependiera de la dirección del espín.) Si ahora permitimos que las direcciones 1 y 2 se aproximen, la amplitud total en la ecuación (4.44) se vuelve cero. El resultado para fermiones es mucho más simple que para bosones. Simplemente no es posible de ninguna manera que dos fermiones —como dos electrones— lleguen exactamente al mismo estado. Nunca encontrarán dos electrones en la misma posición con sus espines en la misma dirección. No es posible para dos electrones tener el mismo momentum y la misma dirección del espín. Si están en el mismo lugar o con el mismo estado de movimiento, la única posibilidad es que deben tener espines opuestos entre sí.

¿Cuáles son las consecuencias de esto? Hay una cantidad de efectos muy notables que son consecuencia del hecho de que dos fermiones no pueden estar en el mismo estado. En realidad, casi todas las particularidades del mundo material giran alrededor de este hecho maravilloso. La variedad que presenta la tabla periódica es, básicamente, consecuencia de esta regla.

Por supuesto, no podemos decir cómo sería el mundo si se cambiara esta regla, porque es una parte de la estructura completa de la mecánica cuántica, y es imposible decir qué más cambiaría si la regla sobre fermiones fuera diferente. De todos modos, tratemos de ver qué ocurriría si únicamente cambiara esta regla. Primero, podemos demostrar que los átomos serían más o menos los mismos. Comencemos con el átomo de hidrógeno. No se vería afectado en forma

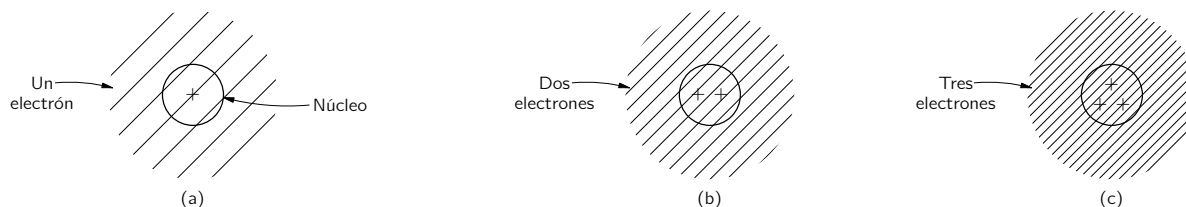


Figura 4-11. Apariencia que tendrían los átomos si los electrones se comportaran como bosones.

apreciable. El protón del núcleo estaría rodeado de una nube electrónica simétrica y esférica, como se muestra en la figura 4-11(a). Como dijimos en el capítulo 2, el electrón es atraído hacia el centro, pero el principio de incertidumbre requiere que exista un balance entre las concentraciones en el espacio y en momentum. Dicho balance significa que debe haber cierta energía y cierta anchura en la distribución electrónica, que determinan las dimensiones características del átomo de hidrógeno.

Ahora supongamos que tenemos un núcleo con dos unidades de carga, como el núcleo de helio. Este núcleo atraerá dos electrones, y si fueran bosones, se agruparían —excepto por su repulsión eléctrica— tan cerca al núcleo como fuera posible. El átomo de helio se podría parecer a lo que muestra la parte (b) de la figura. Análogamente, un átomo de litio, que tiene un núcleo de carga triple, tendría una distribución electrónica como la que muestra la parte (c) de la figura 4-11. Todos los átomos tendrían más o menos la misma apariencia —una pequeña bola redonda con todos los electrones asentados cerca del núcleo, nada direccional y nada complicado—.

No obstante, debido a que los electrones son fermiones, la situación real es bastante diferente. Para el átomo de hidrógeno, la situación es esencialmente la misma. La única diferencia es que el electrón tiene un espín que indicamos con la flecha en la figura 4-12(a). En el caso del átomo de helio, sin embargo, no podemos poner dos electrones uno sobre otro. Un momento, esto es cierto sólo si sus espines son iguales. Dos electrones *pueden* ocupar el mismo estado si sus espines son opuestos. Por lo tanto, el átomo de helio tampoco cambia mucho. Se vería como se muestra en la parte (b) de la figura 4-12. Para el litio, en cambio, la situación es completamente diferente. ¿Dónde podemos poner al tercer electrón? El tercer electrón no puede ir sobre los otros dos, porque ambas direcciones del espín están ocupadas. (Recuerden que para un electrón, o cualquier otra partícula de espín $1/2$, sólo hay dos direcciones posibles para el espín.) El tercer electrón no puede ir cerca del lugar ocupado por los otros dos, por lo que debe quedar en una condición especial, en un estado diferente más alejado del núcleo, como en la parte (c) de la figura. (Estamos hablando un poco a la ligera aquí, porque en realidad los tres electrones son idénticos; como no podemos distinguir cuál es cuál, nuestra descripción es sólo aproximada.)

Ahora podemos empezar a ver por qué diferentes átomos tienen propiedades químicas distintas. Como el tercer electrón del litio está más lejos, está unido de manera relativamente más suelta. Es mucho más fácil remover un electrón del litio que del helio. (Experimentalmente, se necesita 25 voltios para ionizar al helio y sólo 5 voltios para ionizar al litio.) Esto explica la valencia del átomo de litio. Las propiedades direccionales de la valencia tienen que ver con el patrón que forman las ondas del electrón externo, lo cual no explicaremos por el momento. Pero ya podemos observar la importancia del llamado *principio de exclusión* —el cual establece que no se pueden encontrar dos electrones en exactamente el mismo estado (incluyendo el espín)—.

El principio de exclusión también es responsable por la estabilidad de la materia a gran escala. Antes explicamos que los átomos individuales que componen la materia no colapsan a causa del principio de incertidumbre; pero esto no explica por qué no podemos comprimir dos átomos de hidrógeno tanto como

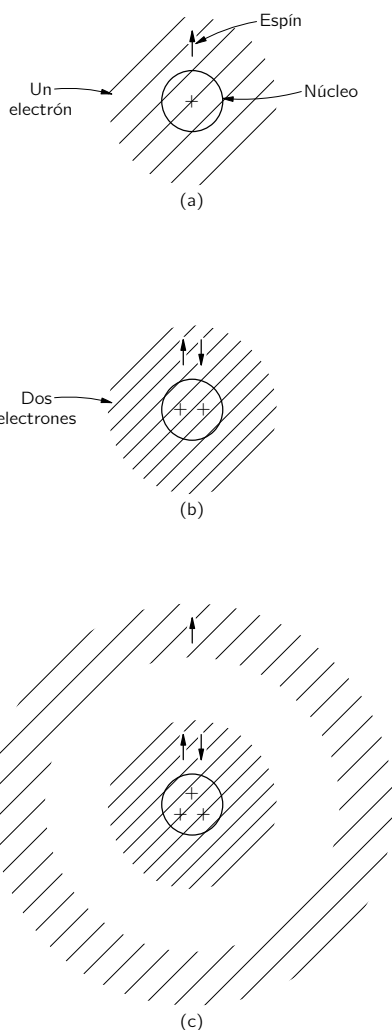


Figura 4-12. Configuraciones atómicas de los electrones reales, de tipo fermión y con espín un medio.

queramos —por qué todos los protones no se juntan con una gran nube de electrones alrededor de ellos—. La respuesta es, naturalmente, que como no más de dos electrones —con espines opuestos— pueden estar en casi el mismo lugar, los átomos de hidrógeno se deben mantener alejados entre sí. Así, la estabilidad de la materia a gran escala es, en realidad, una consecuencia de la naturaleza fermiónica de los electrones.

Es claro que, si los electrones externos de dos átomos tienen espines en direcciones opuestas, éstos pueden acercarse entre sí. Ésta es, en realidad, la manera en que ocurren los enlaces químicos. Resulta que dos átomos juntos tendrán, por lo general, la menor energía posible si hay un electrón entre ellos. Es algo así como una atracción eléctrica de los dos núcleos positivos hacia el electrón en medio. Es posible poner dos electrones más o menos entre dos núcleos, si sus espines son opuestos, y la unión química más fuerte ocurre de esta manera. No hay unión más fuerte, porque el principio de exclusión no permite que haya más de dos electrones entre los átomos. Esperamos que la molécula de hidrógeno sea, más o menos, como se muestra en la figura 4-13.

Queremos mencionar una consecuencia más del principio de exclusión. Ustedes recuerdan que si los dos electrones del átomo de helio van a estar cerca del núcleo, sus espines serán necesariamente opuestos. Supongamos ahora que hacemos un arreglo para tener dos electrones con el mismo espín —como podríamos hacer usando un campo magnético fantásticamente fuerte que tratara de alinear los espines en la misma dirección—. Pero entonces los dos electrones no podrían ocupar el mismo estado en el espacio. Uno de ellos tendría que asumir una posición geométrica diferente, como se indica en la figura 4-14. El electrón que está más lejos del núcleo, está unido a él con menor energía. La energía del átomo completo es, por lo tanto, mucho más alta. En otras palabras, cuando los dos espines son opuestos, hay una atracción total más fuerte.

En consecuencia, hay una enorme fuerza aparente que trata de unir a los espines opuestos cuando dos electrones están muy cerca. Si dos electrones tratan de ir al mismo lugar, hay una tendencia muy fuerte de los espines para alinearse en direcciones opuestas. Esta fuerza aparente, que trata de orientar los dos espines opuestos, es mucho mayor que la fuerza pequeñísima entre los dos momentos magnéticos de los electrones. Recuerden que, cuando hablamos de ferromagnetismo, estaba el misterio de por qué los electrones de átomos diferentes tienen una fuerte tendencia a alinearse paralelamente. Aunque aún no hay una explicación cuantitativa, se cree que esto ocurre porque los electrones alrededor del centro de un átomo interactúan, a través del principio de exclusión, con los electrones externos, que se han liberado para vagar a través del cristal. Esta interacción hace que los espines de los electrones libres y de los electrones internos tomen direcciones opuestas. Pero los electrones libres y los electrones atómicos internos sólo pueden ser opuestos si todos los electrones internos tienen la misma dirección de espín, como se indica en la figura 4-15. Parece probable que sea el efecto del principio de exclusión, actuando indirectamente a través de los electrones libres, lo que provoca las intensas fuerzas de alineación responsables del ferromagnetismo.

Mencionaremos un ejemplo más de la influencia del principio de exclusión. Antes hemos dicho que las fuerzas nucleares son iguales entre neutrón y protón, entre protón y protón, y entre neutrón y neutrón. ¿Por qué un protón y un neutrón se pueden unir para formar un núcleo de deuterio, mientras que no hay ningún núcleo con sólo dos protones o sólo dos neutrones? El deuterio está, de hecho, unido por una energía de unos 2.2 millones de voltios, pero no hay una unión correspondiente entre un par de protones para formar un isótopo del helio con peso atómico 2. Tales núcleos no existen. La combinación de dos protones no forma un estado unido.

La respuesta es el resultado de dos efectos: primero, el principio de exclusión; y segundo, el hecho de que las fuerzas nucleares son de alguna manera sensibles a la dirección del espín. La fuerza entre un neutrón y un protón es atractiva y algo más intensa cuando los espines son paralelos que cuando son opuestos. Sucede que estas fuerzas son suficientemente diferentes que un deuterio sólo se puede

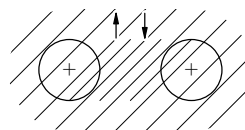


Figura 4-13. La molécula de hidrógeno.

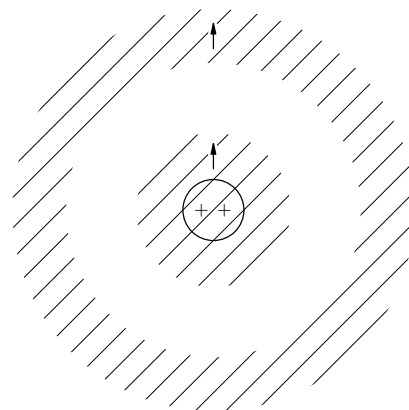


Figura 4-14. Helio con un electrón en un estado de energía más alta.

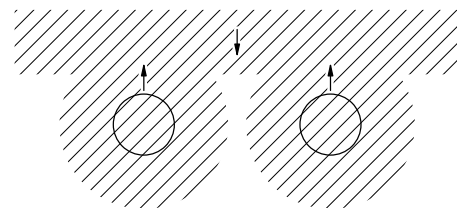


Figura 4-15. El posible mecanismo en un cristal ferromagnético; el electrón de conducción es antiparalelo a los electrones internos no apareados.

formar si el neutrón y el protón tienen espines paralelos; cuando sus espines son opuestos, la atracción no es suficientemente fuerte como para unirlos. Puesto que los espines del neutrón y del protón son de un medio y están en la misma dirección, el deuterio tiene espín uno. Sabemos, sin embargo, que dos protones no pueden colocarse uno sobre otro si sus espines son paralelos. Si no fuera por el principio de exclusión, dos protones podrían unirse, pero como éstos no pueden estar en el mismo lugar con los espines en la misma dirección, el núcleo de He^2 no existe. Los protones podrían estar juntos con sus espines opuestos, pero entonces no habría una unión suficiente para hacer un núcleo estable, porque la fuerza nuclear para espines opuestos es demasiado débil como para unir un par de nucleones. La fuerza atractiva entre neutrones y protones de espines opuestos se puede observar mediante experimentos de dispersión. Experimentos de dispersión similares con dos protones de espines paralelos muestran que sí existe la atracción correspondiente. Así que es el principio de exclusión lo que nos ayuda a explicar por qué el deuterio puede existir y el He^2 no.

Espín uno

5-1 Filtrando átomos con un aparato de Stern-Gerlach

En este capítulo comenzamos propiamente la mecánica cuántica —en el sentido de que vamos a describir un fenómeno cuántico en forma completamente cuántica—. No daremos ninguna excusa ni intentaremos encontrar conexiones con la mecánica clásica. Queremos hablar de algo nuevo en un lenguaje nuevo. La situación específica que vamos a describir es la llamada cuantización del momento angular, para una partícula de *espín uno*. Pero no usaremos palabras como “momento angular” u otros conceptos de la mecánica clásica. Hemos elegido este ejemplo porque es relativamente simple, aunque no el más simple posible. Es suficientemente complicado para poder servir de prototipo, el cual puede generalizarse para describir todos los fenómenos de la mecánica cuántica. Así, aunque estamos tratando un ejemplo particular, todas las leyes que mencionemos se pueden generalizar inmediatamente, y daremos las generalizaciones para que vean las características universales de una descripción cuántica. Comenzaremos con la división de un haz de átomos en tres haces separados en un experimento de Stern-Gerlach.

Recuerdan que si tenemos un campo magnético inhomogéneo formado por un imán con uno de sus polos en punta y enviamos un haz a través del aparato, el haz de partículas se puede dividir en cierto número de haces —el número depende del tipo específico de átomo y de su estado—. Escogeremos el caso de un átomo que produce tres haces, y lo llamaremos partícula de *espín uno*. Ustedes pueden hacer por su cuenta el caso de cinco haces, siete haces, dos haces, etc. —sólo tienen que repetir el procedimiento y donde nosotros tenemos tres términos, ustedes tendrán cinco términos, siete términos, y así—.

Imaginemos el aparato dibujado esquemáticamente en la figura 5-1. Un haz de átomos (o de partículas de cualquier tipo) es colimado por una rejilla y atraviesa un campo no uniforme. Digamos que el haz se mueve en la dirección y , y que el campo magnético y su gradiente están en la dirección z . Entonces, viendo desde adentro, observamos que el haz se divide verticalmente en otros tres, como se muestra en la figura. Ahora, en el extremo de salida del imán podríamos colocar pequeños contadores que contarán la rapidez de llegada de partículas en cualquiera de los tres haces. O podemos bloquear dos de los haces y dejar que el tercero pase.

Supongamos que bloqueamos los dos haces inferiores, y permitimos que pase el haz superior y entre en un segundo aparato de Stern-Gerlach igual al anterior, como se muestra en la figura 5-2. ¿Qué sucede? *No* hay tres haces en el segundo

5-1 Filtrando átomos con un aparato de Stern-Gerlach

5-2 Experimentos con átomos filtrados

5-3 Filtros de Stern-Gerlach en serie

5-4 Estados base

5-5 Amplitudes que interfieren

5-6 La maquinaria de la mecánica cuántica

5-7 Transformando a una base diferente

5-8 Otras situaciones

Revisar

Capítulo 35, volumen II,
Paramagnetismo y resonancias magnéticas.

Por su conveniencia, este capítulo se reproduce en el Apéndice de este volumen.

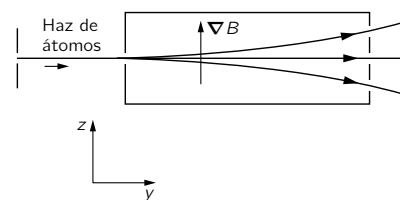


Figura 5-1. En un experimento de Stern-Gerlach, los átomos de espín uno se dividen en tres haces.

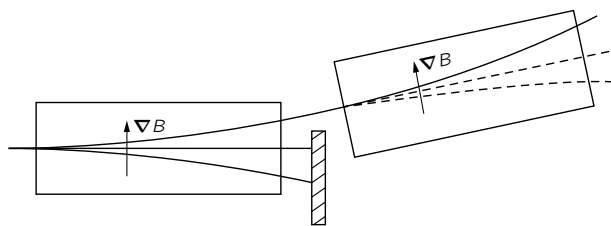


Figura 5-2. Los átomos de uno de los haces se envían a través de un segundo aparato idéntico.

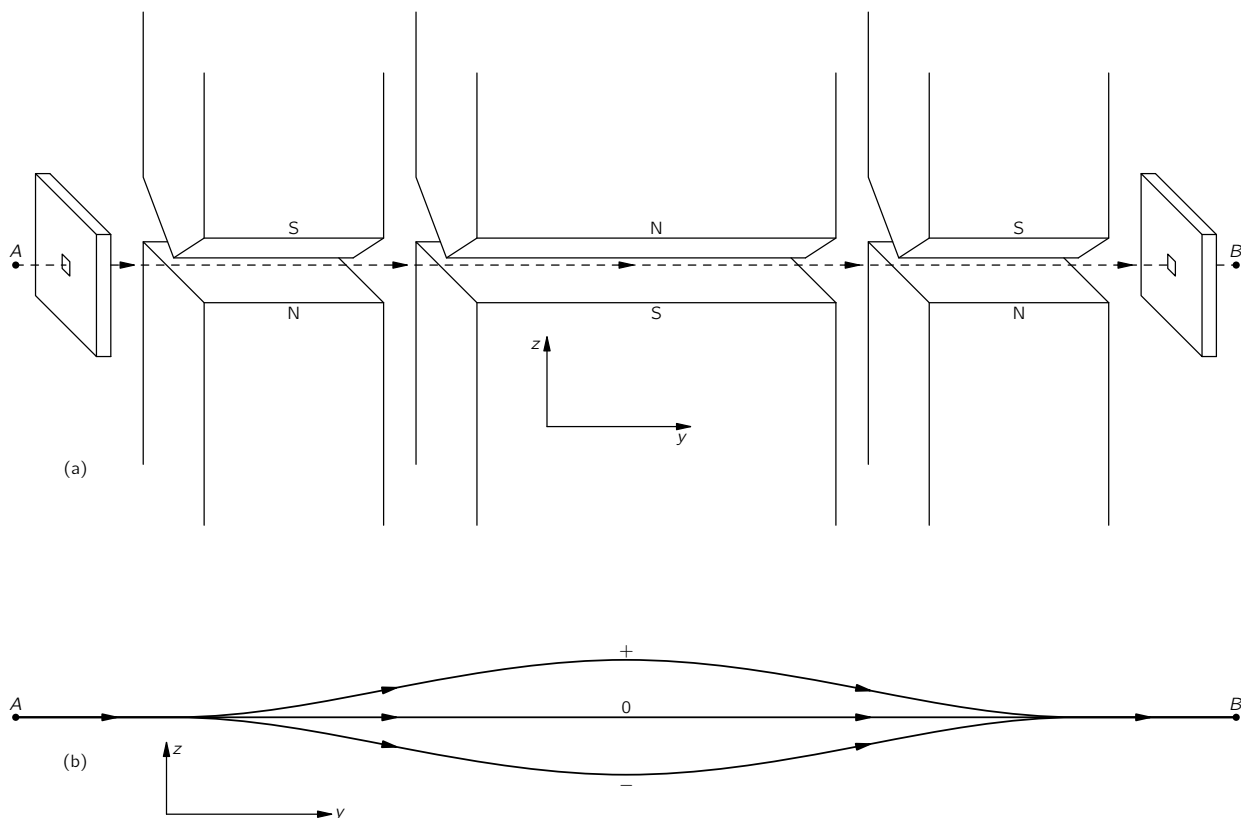


Figura 5-3. (a) Una modificación imaginada de un aparato de Stern-Gerlach. (b) Las trayectorias de átomos con espín uno.

aparato; solamente el haz superior.[†] Esto es lo que esperaríamos si consideramos al segundo aparato como una extensión del primero. Los átomos que están siendo empujados hacia arriba, continúan siendo empujados en el segundo imán.

Pueden ver entonces que el primer aparato ha producido un haz de objetos “purificados” —átomos que son dirigidos hacia arriba en un campo inhomogéneo determinado—. Los átomos, al entrar en el aparato de Stern-Gerlach original, son de tres “variedades”, y cada una de ellas toma una trayectoria diferente. Al filtrar todas excepto una de las variedades, podemos obtener un haz cuyo comportamiento futuro, en un aparato de la misma clase, está determinado y es predecible. Llamaremos a éste un haz *filtrado*, o haz *polarizado*, o un haz cuyos átomos sabemos que se encuentran en un *estado definido*.

Para el resto de nuestra discusión es más conveniente si consideramos un aparato del tipo Stern-Gerlach algo modificado. El aparato parece más complicado en un principio, pero hará que la discusión sea más simple. De todos modos, puesto que solamente son “experimentos mentales”, no cuesta nada complicar el equipo. (A propósito, nadie ha realizado nunca todos los experimentos que describiremos exactamente de esta forma, pero sabemos qué *sucedría* a partir de las leyes de la mecánica cuántica, que están, naturalmente, basadas en otros experimentos similares. Éstos son más difíciles de entender al principio, así que deseamos describir algunos experimentos idealizados —pero posibles—.)

La figura 5-3(a) muestra un dibujo del “aparato modificado de Stern-Gerlach” que nos gustaría utilizar. Éste consta de una secuencia de tres imanes con gradientes grandes. El primero (a la izquierda) es el imán de Stern-Gerlach usual, y divide al haz de partículas de espín uno que entra en tres haces separados. El segundo imán tiene la misma sección transversal que el primero, pero es

[†] Estamos suponiendo que los ángulos de deflexión son muy pequeños.

dos veces más largo, *además* la polaridad de su campo magnético es opuesta al campo del imán 1. El segundo imán empuja en la dirección opuesta a los imanes atómicos y dobla su camino de regreso hacia el eje, como se muestra en las trayectorias dibujadas en la parte inferior de la figura 5-3. El tercer imán es igual al primero, y reúne de nuevo los tres haces, para que pasen por el agujero de salida alineados con el eje. Finalmente, nos gustaría imaginar que frente al agujero en *A* hay un mecanismo que hace posible que los átomos partan del reposo, y que después del agujero de salida en *B* hay un mecanismo de freno que hace que los átomos vuelvan al reposo en *B*. Esto no es necesario, pero significa que en nuestro análisis no tendremos que preocuparnos por incluir cualquier efecto del movimiento cuando los átomos salen, y podemos concentrarnos en aquellos asuntos que sólo tienen relación con el espín. Todo el propósito del aparato “mejorado” es solamente llevar todas las partículas al mismo lugar, y con velocidad cero.

Ahora bien, si queremos realizar un experimento como el de la figura 5-2, primero podemos obtener un haz filtrado, colocando una placa en mitad del aparato que bloquee dos de los haces, como vemos en la figura 5-4. Si ahora hacemos que los átomos polarizados pasen a través de un segundo aparato idéntico, todos los átomos seguirán el camino superior, como se puede verificar colocando placas semejantes en el camino de los distintos haces en el segundo filtro *S* y viendo si pasan partículas.

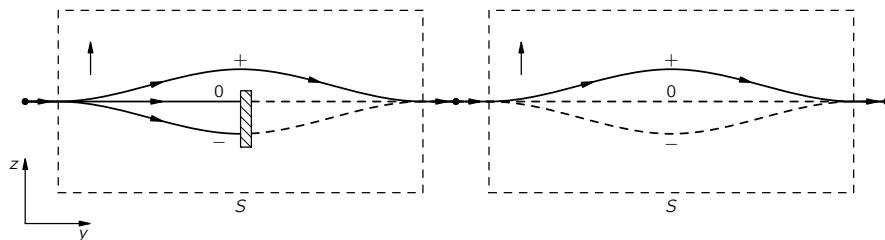


Figura 5-4. El aparato de Stern-Gerlach “mejorado” como un filtro.

Supongamos que llamamos al primer aparato *S*. (Vamos a considerar toda clase de combinaciones, y necesitamos etiquetas para mantener las cosas en orden.) Diremos que los átomos que sigan el camino superior en *S* se encuentran en el “estado más respecto a *S*”; los que tomen el camino de en medio están en el “estado cero respecto a *S*” y aquellos que sigan el camino inferior se encuentran en el “estado menos respecto a *S*”. (En lenguaje más común diríamos que la componente *z* del momento angular es $+1\hbar$, 0 y $-1\hbar$, pero ahora no estamos usando dicho lenguaje.) Ahora, en la figura 5-4 el segundo aparato está orientado exactamente igual que el primero, así que todos los átomos filtrados tomarán el camino superior. O bien, si bloqueamos los haces superior e inferior en el primer aparato, dejando pasar sólo al estado cero, todos los átomos filtrados irán por el camino de en medio del segundo aparato. Y si no hubiésemos bloqueado al haz inferior en el primer aparato, sólo estaría el haz inferior en el segundo. Podemos decir que, en cada caso, nuestro primer aparato ha producido un haz filtrado en un estado *puro* respecto a *S* (+, 0 o -), y podemos comprobar qué estado está presente haciendo pasar a los átomos por un segundo aparato idéntico.

Podemos construir nuestro segundo aparato de tal forma que sólo transmita átomos de un estado determinado —colocando placas dentro como hicimos con el primero— y podemos comprobar el estado del haz que entra simplemente viendo si sale algo al final. Por ejemplo, si bloqueamos los dos caminos más bajos en el segundo aparato, 100 % de los átomos aún pasarán; pero si bloqueamos el camino superior, no pasará ninguno.

Para hacer más simple esta discusión, vamos a inventar un símbolo abreviado para representar a uno de nuestros aparatos de Stern-Gerlach mejorados. Decimos que el símbolo

$$\left\{ \begin{array}{c} + \\ 0 \\ - \end{array} \right\}_S \quad (5.1)$$

representa a un aparato completo. (Éste *no* es un símbolo que encontrarán usado alguna vez en la mecánica cuántica; sólo lo hemos inventado para este capítulo. Simplemente representa una imagen abreviada del aparato de la figura 5-3.) Como deseamos usar varios aparatos al mismo tiempo, y con diversas orientaciones, identificaremos a cada uno de ellos con una letra debajo. El símbolo en (5.1) se refiere al aparato S . Cuando bloqueemos uno o más de los haces adentro, lo indicamos con barras verticales que nos dirán qué haz está bloqueado, de la siguiente manera:

$$\left\{ \begin{array}{c} + \\ 0 \\ - \end{array} \right\}_S \quad (5.2)$$

Las combinaciones posibles que vamos a usar se muestran en la figura 5-5.

Si tenemos dos filtros seguidos (como en la figura 5-4), colocamos los dos símbolos juntos, de esta forma:

$$\left\{ \begin{array}{c} + \\ 0 \\ - \end{array} \right\}_S \quad \left\{ \begin{array}{c} + \\ 0 \\ - \end{array} \right\}_S \quad (5.3)$$

Para este arreglo, todo lo que pasa a través del primero también pasa a través del segundo. En realidad, aunque bloqueemos los canales “cero” y “menos” del segundo aparato, de modo que tengamos

$$\left\{ \begin{array}{c} + \\ 0 \\ - \end{array} \right\}_S \quad \left\{ \begin{array}{c} + \\ 0 \\ - \end{array} \right\}_S, \quad (5.4)$$

aún seguimos obteniendo 100 % de transmisión a través del segundo aparato. Por otra parte, si tenemos

$$\left\{ \begin{array}{c} + \\ 0 \\ - \end{array} \right\}_S \quad \left\{ \begin{array}{c} + \\ 0 \\ - \end{array} \right\}_S, \quad (5.5)$$

nada sale al final. Del mismo modo,

$$\left\{ \begin{array}{c} + \\ 0 \\ - \end{array} \right\}_S \quad \left\{ \begin{array}{c} + \\ 0 \\ - \end{array} \right\}_S \quad (5.6)$$

no saldría nada. Por otra parte,

$$\left\{ \begin{array}{c} + \\ 0 \\ - \end{array} \right\}_S \quad \left\{ \begin{array}{c} + \\ 0 \\ - \end{array} \right\}_S \quad (5.7)$$

sería exactamente igual a

$$\left\{ \begin{array}{c} + \\ 0 \\ - \end{array} \right\}_S$$

por sí mismo.

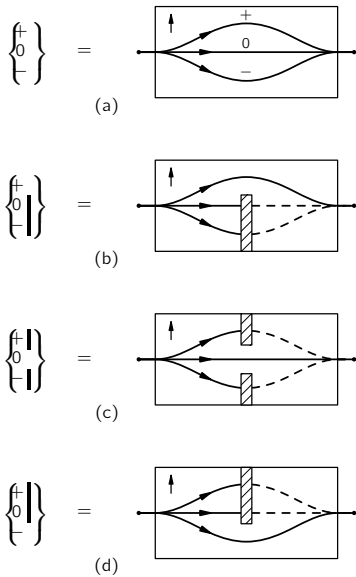


Figura 5-5. Símbolos especiales abreviados para filtros de tipo Stern-Gerlach.

En este momento queremos describir estos experimentos de manera cuántica. Diremos que un átomo se encuentra en el estado $(+S)$ si ha atravesado el aparato de la figura 5-5(b), que se encuentra en el estado $(0S)$ si ha atravesado (c), y en el estado $(-S)$ si ha atravesado (d).[†] Entonces definimos a $\langle b|a \rangle$ como la *amplitud* para que un átomo que se encuentra en el estado a llegue al estado b al atravesar el aparato. Podemos decir: $\langle b|a \rangle$ es la amplitud para que un átomo *en* el estado a *cambie* al estado b . El experimento (5.4) nos da que

$$\langle +S|+S \rangle = 1,$$

mientras que (5.5) nos da

$$\langle -S|+S \rangle = 0.$$

Del mismo modo, el resultado de (5.6) es

$$\langle +S|-S \rangle = 0,$$

y el de (5.7) es

$$\langle -S|-S \rangle = 1.$$

Siempre y cuando tratemos solamente con estados “puros” —esto es, sólo tenemos un canal abierto— hay nueve amplitudes parecidas a éstas, y podemos escribirlas en una tabla:

		desde			
		$+S$	$0S$	$-S$	
hacia	$+S$	1	0	0	(5.8)
	$0S$	0	1	0	
	$-S$	0	0	1	

Este arreglo de nueve números —llamado *matriz*— resume los fenómenos que hemos estado describiendo.

5-2 Experimentos con átomos filtrados

Aquí surge la gran pregunta: ¿qué ocurre si el segundo aparato se gira en un ángulo diferente, de modo que el eje de su campo ya no sea paralelo al primero? Podría no sólo ser girado, sino que también podría apuntar en una dirección distinta —por ejemplo, podría desviar al haz en un ángulo de 90° respecto a la dirección original—. Para tomarlo con calma en un principio, pensemos primero en un arreglo en el cual el segundo experimento de Stern-Gerlach está inclinado en un ángulo α respecto al eje y , como se muestra en la figura 5-6. Llamaremos T al segundo aparato. Supongamos que ahora montamos el siguiente experimento:

$$\left\{ \begin{array}{c} + \\ 0 \\ - \end{array} \right\}_S \quad \left\{ \begin{array}{c} + \\ 0 \\ - \end{array} \right\}_T,$$

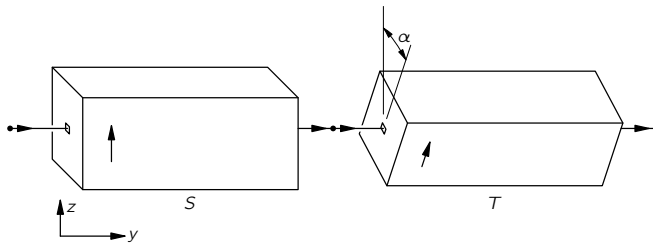


Figura 5-6. Dos filtros de tipo Stern-Gerlach en serie; el segundo está girado un ángulo α respecto al primero.

o el experimento:

[†] Léase: $(+S)$ = “ S -más”, $(0S)$ = “ S -cero” y $(-S)$ = “ S -menos”.

$$\left\{ \begin{array}{c} + \\ 0 \\ - \end{array} \right\}_S \quad \left\{ \begin{array}{c} + \\ 0 \\ - \end{array} \right\}_T.$$

¿Qué ocurre al final en estos casos?

La respuesta es: si los átomos se encuentran en un estado definido respecto a S , *no* se encuentran en el mismo estado respecto a T —un estado $(+S)$ *no* es un estado $(+T)$ —. Hay, sin embargo, una cierta *amplitud* para encontrar al átomo en un estado $(+T)$ —o en un estado $(-T)$, o en uno $(0T)$ —.

En otras palabras, aunque hemos tenido mucho cuidado en asegurarnos de tener a los átomos en una condición definida, la realidad es que, si tienen que atravesar un aparato que está orientado en un ángulo diferente, tienen, por así decirlo, que “reorientarse” —cosa que hacen, no lo olviden, por suerte—. Podríamos hacer que solamente una partícula atravesase en cada momento, y después podríamos preguntarnos: ¿cuál es la probabilidad para que pase a través del aparato? Algunos de los átomos que atraviesen S terminarán en un estado $(+T)$, algunos lo harán en $(-T)$ y otros en un estado $(0T)$ —todos con diferentes probabilidades—. Estas probabilidades se pueden calcular con el cuadrado de las normas de las amplitudes complejas; lo que queremos es algún método matemático, o una descripción cuántica, para estas amplitudes. Necesitamos conocer varias cantidades como

$$\langle -T | +S \rangle,$$

con lo que nos referimos a la amplitud para que un átomo que se encuentra inicialmente en el estado $(+S)$, pueda pasar a la condición $(-T)$ (que *no* es cero, a menos que T y S estén alineados paralelamente). Hay otras amplitudes como

$$\langle +T | 0S \rangle, \quad \text{o} \quad \langle 0T | -S \rangle, \quad \text{etc.}$$

Hay, en realidad, nueve amplitudes de este tipo —otra matriz— que una teoría de partículas debería decirnos cómo calcular. Igual que $F = ma$ nos dice cómo calcular lo que sucede a una partícula clásica en cualquier circunstancia, las leyes de la mecánica cuántica nos permiten calcular la amplitud para que una partícula pase a través de un aparato determinado. El problema principal es, entonces, ser capaces de calcular —para cualquier ángulo de inclinación α dado, o de hecho para cualquier orientación— las nueve amplitudes:

$$\begin{array}{lll} \langle +T | +S \rangle, & \langle +T | 0S \rangle, & \langle +T | -S \rangle, \\ \langle 0T | +S \rangle, & \langle 0T | 0S \rangle, & \langle 0T | -S \rangle, \\ \langle -T | +S \rangle, & \langle -T | 0S \rangle, & \langle -T | -S \rangle. \end{array} \quad (5.9)$$

Ya podemos notar algunas relaciones entre estas amplitudes. En primer lugar, de acuerdo con nuestras definiciones, el cuadrado de la norma

$$|\langle +T | +S \rangle|^2$$

es la *probabilidad* para que un átomo en un estado $(+S)$ llegue al estado $(+T)$. Encontraremos con frecuencia que es más conveniente escribir dichos cuadrados en la forma equivalente

$$\langle +T | +S \rangle \langle +T | +S \rangle^*.$$

En la misma notación, el número

$$\langle 0T | +S \rangle \langle 0T | +S \rangle^*$$

es la probabilidad de que una partícula en el estado $(+S)$ llegue al estado $(0T)$, y

$$\langle -T | +S \rangle \langle -T | +S \rangle^*$$

es la probabilidad para que llegue al estado $(-T)$. Según la forma en que nuestros aparatos están hechos, cada átomo que entre en el aparato T debe estar en *alguno* de los tres estados del aparato T —no hay otro lugar a donde un tipo determinado de átomo pueda ir—. Así que la suma de las tres probabilidades que acabamos de escribir debe ser igual al 100 %. Tenemos la relación

$$\langle +T | +S \rangle \langle +T | +S \rangle^* + \langle 0T | +S \rangle \langle 0T | +S \rangle^* + \langle -T | +S \rangle \langle -T | +S \rangle^* = 1. \quad (5.10)$$

Hay, por supuesto, otras dos ecuaciones del mismo estilo, que obtenemos si comenzamos con el estado $(0S)$ o con el estado $(-S)$. Pero se pueden obtener fácilmente, por eso continuaremos con algunas otras cuestiones generales.

5-3 Filtros de Stern-Gerlach en serie

Surge una pregunta interesante: supongamos que tenemos átomos filtrados en el estado $(+S)$, y los hacemos pasar por un segundo filtro, digamos al estado $(0T)$, y *después* a través de *otro* filtro $+S$. (Llamaremos al último filtro S' para que lo podamos distinguir del primer filtro S .) ¿Los átomos recuerdan que alguna vez estuvieron en un estado $(+S)$? En otras palabras, tenemos el siguiente experimento:

$$\left\{ \begin{array}{c} + \\ 0 \\ - \end{array} \right\}_S \quad \left\{ \begin{array}{c} + \\ 0 \\ - \end{array} \right\}_T \quad \left\{ \begin{array}{c} + \\ 0 \\ - \end{array} \right\}_{S'}. \quad (5.11)$$

Queremos saber si todos aquellos que atraviesen T , también atravesarán S' . *No lo hacen*. Una vez que han sido filtrados por T , *no recuerdan* de ningún modo que se encontraban en un estado $(+S)$ cuando entraron en T . Recuerden que el segundo aparato S en (5.11) está orientado exactamente igual al primero, por lo que sigue siendo un filtro del tipo S . Los estados filtrados por S' son, naturalmente, también $(+S)$, $(0S)$ y $(-S)$.

El punto importante es: *si el filtro T permite pasar sólo a un haz, la fracción que pasa a través del segundo filtro S depende solamente del arreglo del filtro T , es completamente independiente de lo que le precede*. El hecho de que los mismos átomos salieron una vez de un filtro S , no tiene la más mínima influencia sobre qué harán una vez que hayan sido filtrados nuevamente en un haz puro mediante un aparato T . Entonces, la probabilidad para llegar a estados diferentes es la misma, sin importar qué sucedió antes de que entraran en el aparato T .

Por ejemplo, comparemos el experimento de (5.11) con el siguiente:

$$\left\{ \begin{array}{c} + \\ 0 \\ - \end{array} \right\}_S \quad \left\{ \begin{array}{c} + \\ 0 \\ - \end{array} \right\}_T \quad \left\{ \begin{array}{c} + \\ 0 \\ - \end{array} \right\}_{S'}. \quad (5.12)$$

en el cual sólo se ha cambiado el primer S . Digamos que el ángulo α (entre S y T) es tal que, en el experimento (5.11), un tercio de los átomos que atraviesan T también atraviesan S' . En el experimento (5.12), aunque habrá, en general, un número diferente de átomos que atraviesen T , *la misma fracción de éstos* —es decir, un tercio— también atraviesan S' .

Podemos, de hecho, mostrarles a partir de lo que han aprendido antes, que la fracción de átomos que sale de T y llegan a un S' determinado depende solamente de T y de S' , y no de nada que haya sucedido antes. Comparemos el experimento (5.12) con

$$\left\{ \begin{array}{c} + \\ 0 \\ - \end{array} \right\}_S \quad \left\{ \begin{array}{c} + \\ 0 \\ - \end{array} \right\}_T \quad \left\{ \begin{array}{c} + \\ 0 \\ - \end{array} \right\}_{S'}. \quad (5.13)$$

La amplitud para que un átomo que sale de S , también atraviese T y S' es, para el experimento en (5.12),

$$\langle +S | 0T \rangle \langle 0T | 0S \rangle.$$

La probabilidad correspondiente es

$$|\langle +S | 0T \rangle \langle 0T | 0S \rangle|^2 = |\langle +S | 0T \rangle|^2 |\langle 0T | 0S \rangle|^2.$$

La probabilidad para el experimento (5.13) es

$$|\langle 0S | 0T \rangle \langle 0T | 0S \rangle|^2 = |\langle 0S | 0T \rangle|^2 |\langle 0T | 0S \rangle|^2.$$

El cociente de ambas es

$$\frac{|\langle 0S | 0T \rangle|^2}{|\langle +S | 0T \rangle|^2}$$

que depende solamente de T y S' , y de no de cuál haz $(+S)$, $(0S)$ o $(-S)$ seleccionó S . (Las normas pueden aumentar o disminuir al mismo tiempo, dependiendo de cuántos atraviesan T .) Encontraríamos, por supuesto, el mismo resultado, si comparásemos las probabilidades para que los átomos alcancen los estados más o menos respecto a S' , o el cociente de las probabilidades para que alcancen los estados cero o menos.

De hecho, ya que estos cocientes dependen solamente de a cuál haz se le permite atravesar T , y no de la selección hecha por el primer filtro S , resulta claro que obtendríamos el mismo resultado aunque el último aparato no fuese un filtro S . Si usamos como tercer aparato —que ahora llamaremos R — uno rotado un ángulo arbitrario respecto a T , encontraríamos que un cociente como $|\langle 0R | 0T \rangle|^2 / |\langle +R | 0T \rangle|^2$ es independiente del haz que pasa por el primer filtro S .

5-4 Estados base

Estos resultados ilustran uno de los principios básicos de la mecánica cuántica: cualquier sistema atómico se puede separar, mediante un proceso de filtración, en un cierto conjunto que llamaremos *estados base*, y el comportamiento futuro de los átomos en cualquier estado base determinado dependerá solamente de la naturaleza del estado base —es independiente de cualquier historial previo—.† Los estados base dependen, naturalmente, del filtro usado; por ejemplo, los tres estados $(+T)$, $(0T)$ y $(-T)$ son un conjunto de estados base; los tres estados $(+S)$, $(0S)$ y $(-S)$ son otro. Hay un gran número de posibilidades, cada una tan válida como las otras.

Deberíamos ser cuidadosos y decir que estamos considerando *buenos* filtros, los cuales producen en verdad haces “puros”. Si, por ejemplo, nuestro aparato de Stern-Gerlach no produjera una buena separación de los tres haces, de modo que no podamos separarlos limpiamente con nuestras placas, entonces no podríamos hacer una separación completa en estados base. Podemos decir si tenemos estados base puros, viendo si los haces se pueden separar de nuevo en otro filtro de la misma clase. Si tenemos un estado puro $(+T)$, por ejemplo, todos los átomos pasarán a través de

$$\left\{ \begin{array}{c} + \\ 0 \\ - \end{array} \right\}_T,$$

y ninguno pasará a través de

$$\left\{ \begin{array}{c} + \\ 0 \\ - \end{array} \right\}_T,$$

† No pretendemos que la palabra “estado base” implique algo más que lo dicho aquí. No se debe considerar como algo “básico” en sentido alguno. Estamos usando la palabra base pensando en una *base* para una descripción, algo así como el sentido que se le da cuando se habla de “los números en *base* diez”.

o a través de

$$\left\{ \begin{array}{c} + \\ 0 \\ - \end{array} \right\}_T.$$

Nuestra afirmación sobre los estados base quiere decir que es *posible* filtrar un estado puro, de modo que no es posible filtrarlo de nuevo mediante un aparato idéntico.

También debemos señalar que, lo que estamos diciendo, es estrictamente verdadero sólo en situaciones un tanto idealizadas. En cualquier aparato de Stern-Gerlach real, tendríamos que preocuparnos por la difracción causada por las rejillas, que podría ocasionar que algunos átomos fuesen a estados correspondientes a ángulos diferentes, o si los haces contienen átomos con diferentes excitaciones en sus estados internos, y otras cosas más. Hemos idealizado la situación para hablar solamente de los estados que son divididos en un campo magnético; estamos ignorando aspectos relacionados con la posición, el momentum, las excitaciones internas, y cuestiones parecidas. En general, también se necesitarían considerar estados base que son agrupados respecto a todas esas cosas. Pero, para mantener los conceptos simples, sólo estamos considerando un conjunto de tres estados, que es suficiente para el tratamiento exacto de la situación idealizada, en la que los átomos no son desgarrados cuando atraviesan el aparato, o maltratados de algún otro modo, y llegan al reposo cuando salen del aparato.

Habrán notado que siempre comenzamos nuestros experimentos mentales tomando un filtro con sólo un canal abierto, así que comenzamos con algún estado base definido. Hacemos esto porque los átomos salen de un horno en varios estados determinados al azar por sucesos accidentales dentro del horno. (Produce lo que se llama un haz “no polarizado”.) Este azar involucra probabilidades del tipo “clásico” —como al lanzar una moneda— que son diferentes de las probabilidades cuánticas que nos interesan. Tratar con un haz no polarizado provocaría complicaciones adicionales, que será mejor evitar hasta que comprendamos el comportamiento de haces polarizados. Así que no traten de considerar en este momento qué ocurre si el *primer* aparato deja pasar más de un haz. (Les diremos cómo pueden tratar tales casos al final del capítulo.)

Ahora regresemos y veamos qué sucede cuando vamos desde un estado base para un filtro, a otro estado base para otro filtro diferente. Supongamos que empezamos de nuevo con

$$\left\{ \begin{array}{c} + \\ 0 \\ - \end{array} \right\}_S \quad \left\{ \begin{array}{c} + \\ 0 \\ - \end{array} \right\}_T.$$

Los átomos que salen de T están en el estado base $(0T)$ y no recuerdan que estuvieron alguna vez en el estado $(+S)$. Algunas personas dirían que al filtrar por T hemos “perdido la información” acerca del estado previo $(+S)$, porque hemos “alterado” a los átomos cuando los separamos en tres haces en el aparato T . Pero esto no es verdad. La información anterior no se ha perdido por la *separación* en tres haces, sino por las *placas bloqueadoras* que hemos colocado —como pueden ver en el siguiente conjunto de experimentos—.

Empezamos con un filtro $+S$, y llamaremos N al número de átomos que lo atraviesan. Si seguimos con un filtro $0T$, el número de átomos que saldrá es una fracción del número original, digamos αN . Si después ponemos otro filtro $+S$, solamente una fracción β de estos átomos llegará el final. Podemos indicar esto en la siguiente forma:

$$\left\{ \begin{array}{c} + \\ 0 \\ - \end{array} \right\}_S \xrightarrow{N} \left\{ \begin{array}{c} + \\ 0 \\ - \end{array} \right\}_T \xrightarrow{\alpha N} \left\{ \begin{array}{c} + \\ 0 \\ - \end{array} \right\}_{S'} \xrightarrow{\beta \alpha N} \quad (5.14)$$

Si nuestro tercer aparato S' seleccionase un estado diferente, digamos el estado $(0S)$, una fracción diferente, digamos γ , lo atravesaría.[†] Tendríamos

$$\left\{ \begin{array}{c} + \\ 0 \\ - \end{array} \right\}_S \xrightarrow{N} \left\{ \begin{array}{c} + \\ 0 \\ - \end{array} \right\}_T \xrightarrow{\alpha N} \left\{ \begin{array}{c} + \\ 0 \\ - \end{array} \right\}_{S'} \xrightarrow{\gamma \alpha N} \quad (5.15)$$

Supongamos ahora que repetimos estos dos experimentos, pero quitamos todas las placas de T . Encontraríamos los sorprendentes resultados que siguen:

$$\left\{ \begin{array}{c} + \\ 0 \\ - \end{array} \right\}_S \xrightarrow{N} \left\{ \begin{array}{c} + \\ 0 \\ - \end{array} \right\}_T \xrightarrow{N} \left\{ \begin{array}{c} + \\ 0 \\ - \end{array} \right\}_{S'} \xrightarrow{N} \quad (5.16)$$

$$\left\{ \begin{array}{c} + \\ 0 \\ - \end{array} \right\}_S \xrightarrow{N} \left\{ \begin{array}{c} + \\ 0 \\ - \end{array} \right\}_T \xrightarrow{N} \left\{ \begin{array}{c} + \\ 0 \\ - \end{array} \right\}_{S'} \xrightarrow{0} \quad (5.17)$$

¡*Todos* los átomos atraviesan S' en el primer caso, pero *ninguno* en el segundo caso! Ésta es una de las grandes leyes de la mecánica cuántica. Que la naturaleza trabaje de esta manera no es evidente por sí mismo, pero los resultados que hemos dado corresponden, en nuestra situación idealizada, al comportamiento cuántico observado en innumerables experimentos.

5-5 Amplitudes que interfieren

¿Cómo puede ser que al ir de (5.15) a (5.17) —*abriendo más canales*— dejamos pasar *menos* átomos? Éste es el antiguo y profundo misterio de la mecánica cuántica —la interferencia de amplitudes—. Es algo semejante a lo que vimos en el experimento de interferencia a través de dos rejillas con electrones. Vimos que podíamos obtener menos electrones en algunos lugares con las dos rejillas abiertas, que los que obteníamos con una sola rejilla abierta. Funciona cuantitativamente así. Podemos escribir la amplitud para que un átomo atravesase T y S' en el aparato de (5.17) como la suma de tres amplitudes, una para cada uno de los haces en T ; esta suma es igual a cero:

$$\langle 0S | +T \rangle \langle +T | +S \rangle + \langle 0S | 0T \rangle \langle 0T | +S \rangle + \langle 0S | -T \rangle \langle -T | +S \rangle = 0. \quad (5.18)$$

Ninguna de las tres amplitudes individualmente es cero —por ejemplo, el cuadrado de la norma de la segunda amplitud es $\gamma\alpha$, vean (5.15)— pero la *suma es cero*. Tendríamos la misma respuesta si S' fuera arreglado para seleccionar el estado $(-S)$. Sin embargo, en el arreglo de (5.16), la respuesta es diferente. Si llamamos a a la amplitud para pasar a través de T y S' , en este caso tenemos[‡]

$$a = \langle +S | +T \rangle \langle +T | +S \rangle + \langle +S | 0T \rangle \langle 0T | +S \rangle + \langle +S | -T \rangle \langle -T | +S \rangle = 1. \quad (5.19)$$

En el experimento (5.16) el haz ha sido dividido y recombinado. Humpty Dumpty se ha reconstruido.[¶] La información acerca del estado original $(+S)$ se mantiene —es como si el aparato T no se encontrase ahí en lo absoluto—.

[†] En nuestra notación anterior $\alpha = |\langle 0T | +S \rangle|^2$, $\beta = |\langle +S | 0T \rangle|^2$ y $\gamma = |\langle 0S | 0T \rangle|^2$.

[‡] En realidad no podemos concluir a partir del experimento que $a = 1$, sino solamente que $|a|^2 = 1$, por lo que a debería ser $e^{i\delta}$, pero se puede demostrar que escoger $\delta = 0$ no representa una pérdida de generalidad.

[¶] N. del T.: el autor se refiere a una canción infantil sobre un huevo personificado llamado Humpty Dumpty, la cual dice: “Humpty Dumpty se sentó sobre un muro, Humpty Dumpty se desplomó muy duro. Acudieron todos los hombres y caballos del rey, pero Humpty Dumpty no pudo ser reconstruido”.

Esto es verdad, sin importar lo que se coloque después del aparato T “abierto por completo”. Podríamos colocar adelante un filtro R —un filtro en cualquier ángulo raro— o lo que queramos. La respuesta siempre será la misma, como si los átomos se tomaran directamente del primer filtro S .

Éste es el principio importante: un filtro T —o cualquier filtro— con placas abiertas no produce ningún cambio. Podríamos enunciar una condición adicional. El filtro abierto por completo no sólo debe transmitir los tres haces, sino que además *no* debe producir alteraciones desiguales en los tres haces. Por ejemplo, no debe tener un campo eléctrico fuerte cerca de un haz y no tenerlo cerca de los otros. La razón es que, aunque esta alteración adicional dejara pasar los átomos a través del filtro, podría cambiar las *fases* de algunas amplitudes. Entonces la interferencia cambiaría, y las amplitudes en las ecuaciones (5.18) y (5.19) serían diferentes. También suponemos que no existen tales alteraciones adicionales.

Volvamos a escribir las ecuaciones (5.18) y (5.19) en una notación mejorada. Diremos que i representa a cualquiera de los tres estados $(+T)$, $(0T)$ o $(-T)$; entonces las ecuaciones se pueden escribir:

$$\sum_{\text{toda } i} \langle 0S | i \rangle \langle i | +S \rangle = 0 \quad (5.20)$$

y

$$\sum_{\text{toda } i} \langle +S | i \rangle \langle i | +S \rangle = 1. \quad (5.21)$$

Del mismo modo, para un experimento donde S' es remplazado por un filtro R completamente arbitrario, tenemos

$$\left\{ \begin{array}{c} + \\ 0 \\ - \end{array} \right\}_S \quad \left\{ \begin{array}{c} + \\ 0 \\ - \end{array} \right\}_T \quad \left\{ \begin{array}{c} + \\ 0 \\ - \end{array} \right\}_R. \quad (5.22)$$

Los resultados siempre serán los mismos, como si el aparato T se hubiese quitado y tuviésemos solamente

$$\left\{ \begin{array}{c} + \\ 0 \\ - \end{array} \right\}_S \quad \left\{ \begin{array}{c} + \\ 0 \\ - \end{array} \right\}_R.$$

O, expresado matemáticamente,

$$\sum_{\text{toda } i} \langle +R | i \rangle \langle i | +S \rangle = \langle +R | +S \rangle. \quad (5.23)$$

Ésta es nuestra ley fundamental, y es verdadera de manera general, siempre que i represente los tres estados de base de cualquier filtro.

Notarán que en el experimento (5.22) no hay una relación espacial entre S y R con T . Además, el razonamiento es el mismo sin importar los estados que éstos seleccionen. Para escribir la ecuación de forma general, sin tener que referirnos a los estados específicos seleccionados por S y R , llamaremos ϕ (“phi”) al estado preparado por el primer filtro (en nuestro ejemplo especial, $+S$) y χ (“chi”) al estado comprobado por el filtro final (en nuestro ejemplo, $+R$). Entonces podemos establecer nuestra ley fundamental de la ecuación (5.23) en la forma

$$\langle \chi | \phi \rangle = \sum_{\text{toda } i} \langle \chi | i \rangle \langle i | \phi \rangle, \quad (5.24)$$

donde i recorre los tres estados base de algún filtro determinado.

Queremos insistir nuevamente en lo que entendemos por estados base. Son como los tres estados que se puede seleccionar por uno de nuestros aparatos de Stern-Gerlach. Una de las condiciones es que si se tiene un estado base, entonces el futuro es independiente del pasado. Otra condición es que si ustedes tienen un

conjunto completo de estados base, entonces la ecuación (5.24) es válida para cualquier conjunto de estados iniciales y finales ϕ y χ . No hay, sin embargo, un *único* conjunto de estados base. Comenzamos considerando estados base *respecto a* un aparato T determinado. Podríamos haber considerado igualmente un *conjunto diferente* de estados base, respecto a un aparato S , o respecto a R , etc.† Por lo general hablamos de estados base “en una cierta representación”.

Otra condición para un conjunto de estados base en cualquier representación determinada, es que éstos sean completamente diferentes. Con esto queremos decir que si tenemos un estado $(+T)$, no existe una amplitud para que éste pase a un estado $(0T)$ o $(-T)$. Si decimos que i y j representan cualesquiera dos estados base de un conjunto determinado, las reglas generales que discutimos con relación a (5.8) son que

$$\langle j | i \rangle = 0$$

para todas las i y j que no sean iguales. Naturalmente, sabemos que

$$\langle i | i \rangle = 1.$$

Estas dos ecuaciones se escriben por lo general en la forma

$$\langle j | i \rangle = \delta_{ji}, \quad (5.25)$$

donde δ_{ji} (la “delta de Kronecker”) es un símbolo definido como cero para $i \neq j$ y uno para $i = j$.

La ecuación (5.25) no es independiente de las otras leyes que hemos mencionado. Sucede que no estamos particularmente interesados en el problema matemático de encontrar el conjunto mínimo de axiomas independientes, que genere todas las leyes como consecuencias.‡ Nos contentamos con tener un conjunto que sea completo y no inconsistente. Podemos, sin embargo, demostrar que las ecuaciones (5.25) y (5.24) no son independientes. Supongamos que ϕ en la ecuación (5.24) representa a uno de los estados base del mismo conjunto como i , digamos el estado j -ésimo; entonces tenemos

$$\langle \chi | j \rangle = \sum_i \langle \chi | i \rangle \langle i | j \rangle.$$

Pero la ecuación (5.25) dice que $\langle i | j \rangle$ es cero a menos que $i = j$, por lo que la suma resulta $\langle \chi | j \rangle$ y obtenemos una identidad, lo que muestra que las dos leyes no son independientes.

Podemos ver que debe haber otra relación entre las amplitudes, si las ecuaciones (5.10) y (5.24) son verdaderas. La ecuación (5.10) es

$$\langle +T | +S \rangle \langle +T | +S \rangle^* + \langle 0T | +S \rangle \langle 0T | +S \rangle^* + \langle -T | +S \rangle \langle -T | +S \rangle^* = 1.$$

Si escribimos la ecuación (5.24), suponiendo que ϕ y χ son el estado $(+S)$, el lado izquierdo es $\langle +S | +S \rangle$, que evidentemente vale 1; así que obtenemos una vez más la ecuación (5.19),

$$\langle +S | +T \rangle \langle +T | +S \rangle + \langle +S | 0T \rangle \langle 0T | +S \rangle + \langle +S | -T \rangle \langle -T | +S \rangle = 1.$$

Estas dos ecuaciones son consistentes (para todas las orientaciones relativas de los aparatos T y S) solamente si

$$\begin{aligned} \langle +S | +T \rangle &= \langle +T | +S \rangle^*, \\ \langle +S | 0T \rangle &= \langle 0T | +S \rangle^*, \\ \langle +S | -T \rangle &= \langle -T | +S \rangle^*. \end{aligned}$$

† De hecho, para sistemas atómicos con tres o más estados base, existen otros tipos de filtros —bastante diferentes al aparato de Stern-Gerlach— que se pueden usar para obtener más posibilidades para el conjunto de estados base (cada conjunto con el mismo número de estados).

‡ ¡La *verdad* redundante no nos molesta!

Y se sigue que para cualesquiera estados ϕ y χ ,

$$\langle \phi | \chi \rangle = \langle \chi | \phi \rangle^*. \quad (5.26)$$

Si esto no fuese verdad, no se “conservaría” la probabilidad, y las partículas se “perderían”.

Antes de continuar, queremos resumir las tres leyes generales importantes acerca de las amplitudes. Éstas son las ecuaciones (5.24), (5.25) y (5.26):

$$\begin{aligned} \text{I} \quad & \langle j | i \rangle = \delta_{ji}, \\ \text{II} \quad & \langle \chi | \phi \rangle = \sum_{\text{toda } i} \langle \chi | i \rangle \langle i | \phi \rangle, \\ \text{III} \quad & \langle \phi | \chi \rangle = \langle \chi | \phi \rangle^*. \end{aligned} \quad (5.27)$$

En estas ecuaciones las i y j se refieren a *todos* los estados base de *alguna* representación, mientras que ϕ y χ representan cualquier estado posible de los átomos. Es importante notar que II es válida solamente si la suma se hace sobre *todos* los estados base del sistema (en nuestro caso, tres: $+T$, $0T$, $-T$). Estas leyes no nos dicen nada acerca de qué debemos escoger como base para nuestro conjunto de estados base. Comenzamos usando un aparato T , que es un experimento de Stern-Gerlach con cierta orientación arbitraria; pero cualquier otra orientación, digamos W , sería igual de buena como ésta. Tendríamos un conjunto diferente de estados para usar como i y j , pero todas las leyes seguirían siendo válidas —no hay un conjunto único—. Uno de los grandes juegos de la mecánica cuántica es aprovecharse de que las cosas se pueden calcular en más de una forma.

5-6 La maquinaria de la mecánica cuántica

Queremos mostrarles por qué estas leyes son útiles. Supongan que tenemos un átomo en una condición dada (con lo cual queremos dar a entender que lo hemos preparado de cierto modo), y queremos saber qué le sucederá en algún experimento. En otras palabras, empezamos con nuestro átomo en el estado ϕ y queremos saber cuáles son las *posibilidades* para que atraviese un aparato que sólo acepta átomos en el estado χ . Las leyes dicen que podemos describir por completo al aparato en términos de tres números complejos $\langle \chi | i \rangle$, las amplitudes para que cada estado base se encuentre en la condición χ ; y que podemos decir qué sucederá si un átomo se coloca dentro del aparato si describimos el estado del átomo dando tres números $\langle i | \phi \rangle$, las amplitudes para que el átomo en su condición inicial se encuentre en cada uno de los tres estados base. Ésta es una idea de gran importancia.

Consideremos otro ejemplo. Piensen en el siguiente problema: empezamos con un aparato S ; luego tenemos un laberinto muy complicado, que podemos llamar A , y después un aparato R —así—:

$$\left\{ \begin{array}{c} + \\ 0 \\ - \end{array} \right\} \left\{ A \right\} \left\{ \begin{array}{c} + \\ 0 \\ - \end{array} \right\}. \quad (5.28)$$

$S \qquad \qquad R$

Por A entendemos cualquier arreglo complicado de aparatos de Stern-Gerlach con placas o medias placas, orientados en cualquier ángulo, con campos eléctricos y magnéticos extraños... prácticamente todo lo que quieran poner. (Es muy agradable realizar experimentos mentales —¡no tienen que pasar por todo el problema de realmente *construir* el aparato!—.) El problema es entonces: ¿Con qué amplitud una partícula que entra en la sección A en un estado $(+S)$, sale de A en el estado $(0R)$, de modo que pueda atravesar el último filtro R ? La notación común para esta amplitud es

$$\langle 0R | A | +S \rangle.$$

Como de costumbre, se tiene que leer de derecha a izquierda (como en hebreo):

$$\langle \text{final} | \text{a través} | \text{inicio} \rangle.$$

Si, por casualidad, A no hace nada —es sólo un canal abierto— entonces escribimos

$$\langle 0R | 1 | +S \rangle = \langle 0R | +S \rangle; \quad (5.29)$$

los dos símbolos son equivalentes. En un problema más general, podríamos remplazar a $(+S)$ por un estado inicial general ϕ y a $(0R)$ por un estado final general χ , y querríamos conocer la amplitud

$$\langle \chi | A | \phi \rangle.$$

Un análisis completo del aparato A tendría que dar la amplitud $\langle \chi | A | \phi \rangle$ para cada par posible de estados ϕ y χ —¡un número infinito de combinaciones!—. ¿Cómo podemos entonces dar una descripción consistente del comportamiento del aparato A ? Podemos hacerlo del siguiente modo. Imaginen que el aparato de (5.28) se modifica para ser

$$\begin{array}{ccccc} \left\{ \begin{array}{c} + \\ 0 \\ - \end{array} \right\} & \left\{ \begin{array}{c} + \\ 0 \\ - \end{array} \right\} & \left\{ A \right\} & \left\{ \begin{array}{c} + \\ 0 \\ - \end{array} \right\} & \left\{ \begin{array}{c} + \\ 0 \\ - \end{array} \right\} \\ S & T & & T & R \end{array} \quad (5.30)$$

Ésta no es realmente una modificación ya que los aparatos T totalmente abiertos no hacen nada. Pero sí nos sugieren cómo podemos analizar el problema. Hay un cierto conjunto de amplitudes $\langle i | +S \rangle$ para que los átomos de S pasen al estado i de T . Después hay otro conjunto de amplitudes para que un estado i (respecto a T) que entra en A salga como un estado j (respecto a T). Y, finalmente, hay una amplitud para que cada estado j pase a través del último filtro en un estado $(0R)$. Para cada camino posible, hay una amplitud de la forma

$$\langle 0R | j \rangle \langle j | A | i \rangle \langle i | +S \rangle,$$

y la amplitud total es la suma de los términos que obtenemos con todas las combinaciones posibles de i y j . La amplitud que queremos es

$$\sum_{ij} \langle 0R | j \rangle \langle j | A | i \rangle \langle i | +S \rangle. \quad (5.31)$$

Si $(0R)$ y $(+S)$ son remplazados por estados generales χ y ϕ , obtenemos el mismo tipo de expresión; por lo que tenemos el resultado general

$$\langle \chi | A | \phi \rangle = \sum_{ij} \langle \chi | j \rangle \langle j | A | i \rangle \langle i | \phi \rangle. \quad (5.32)$$

Observen ahora que el lado derecho de la ecuación (5.32) es realmente “más simple” que el lado izquierdo. El aparato A queda completamente determinado por los *nueve* números $\langle j | A | i \rangle$, que nos dan la respuesta de A respecto a los tres estados base del aparato T . Una vez que conocemos estos nueve números, podemos manejar cualquier par de estados entrantes y salientes χ y ϕ , si los definimos en términos de las tres amplitudes para ir a, o venir desde, cada uno de los tres estados base. El resultado de un experimento se predice usando la ecuación (5.32).

Ésta es entonces la maquinaria de la mecánica cuántica para una partícula de espín uno. Cada *estado* se describe con tres números, que son las amplitudes para encontrarse en cada uno de los estados base que se han seleccionado. Cada aparato se describe con nueve números, que son las amplitudes para ir de un estado base a otro en el aparato. A partir de estos números se puede realizar cualquier cálculo.

Las nueve amplitudes que describen al aparato, por lo general, se escriben como una matriz cuadrada —llamada la matriz $\langle j | A | i \rangle$ —:

$$\begin{array}{c} \text{desde} \\ \begin{array}{ccc} + & 0 & - \\ \text{hacia} & + & 0 & - \end{array} \end{array} \begin{array}{c} \left| \begin{array}{ccc} \langle + | A | + \rangle & \langle + | A | 0 \rangle & \langle + | A | - \rangle \\ \langle 0 | A | + \rangle & \langle 0 | A | 0 \rangle & \langle 0 | A | - \rangle \\ \langle - | A | + \rangle & \langle - | A | 0 \rangle & \langle - | A | - \rangle \end{array} \right| \end{array} \quad (5.33)$$

Las matemáticas de la mecánica cuántica son sólo una extensión de esta idea. Les daremos un ejemplo simple. Supongan que tenemos un aparato C que queremos analizar —es decir, queremos calcular los diversos $\langle j | C | i \rangle$ —. Por ejemplo, podríamos querer saber qué sucede en un experimento como

$$\left\{ \begin{array}{c} + \\ 0 \\ - \end{array} \right\}_S \quad \left\{ C \right\} \quad \left\{ \begin{array}{c} + \\ 0 \\ - \end{array} \right\}_R. \quad (5.34)$$

Pero entonces nos damos cuenta que el aparato C está formado por los dos aparatos A y B en serie —las partículas pasan por A y luego por B — así que podemos escribir simbólicamente

$$\left\{ C \right\} = \left\{ A \right\} \cdot \left\{ B \right\}. \quad (5.35)$$

Podemos llamar al aparato C el “producto” de A y B . Supongan también que ya sabemos cómo analizar las dos partes; por lo que podemos obtener las matrices (respecto a T) de A y de B . Nuestro problema está resuelto entonces. Podemos encontrar fácilmente

$$\langle \chi | C | \phi \rangle$$

para estados cualesquiera de entrada y salida. Primero escribimos

$$\langle \chi | C | \phi \rangle = \sum_k \langle \chi | B | k \rangle \langle k | A | \phi \rangle.$$

¿Ven por qué? (*Sugerencia:* Imaginen que colocan un aparato T entre A y B .) Así, si consideramos el caso especial en que ϕ y χ son también estados base (de T), digamos i y j , tenemos

$$\langle j | C | i \rangle = \sum_k \langle j | B | k \rangle \langle k | A | i \rangle. \quad (5.36)$$

Esta ecuación da la matriz para el aparato “producto” C en términos de las dos matrices de los aparatos A y B . Los matemáticos llaman a la nueva matriz $\langle j | C | i \rangle$ —formada a partir de las dos matrices $\langle j | B | i \rangle$ y $\langle j | A | i \rangle$ de acuerdo con la suma especificada en (5.36)— el “producto” matricial BA de las dos matrices B y A . (Noten que el *orden* es importante, $AB \neq BA$.) Así, podemos decir que la matriz para una sucesión de dos aparatos es el producto matricial de las dos matrices para los dos aparatos (colocando al *primer* aparato a la *derecha* en el producto). Cualquiera que sepa de álgebras de matrices entiende que nos referimos exactamente a la ecuación (5.36).

5-7 Transformando a una base diferente

Queremos dar un último argumento sobre los estados base usados en los cálculos. Supongan que hemos elegido trabajar con una base particular —digamos la base S — y otro compañero decide hacer los mismos cálculos con una base diferente —digamos la base T —. Para mantener todo en orden, llamemos a

nuestros estados base los estados (iS) , donde $i = +, 0, -$. Del mismo modo, podemos llamar a sus estados base (jT) . ¿Cómo podemos comparar nuestro trabajo con el de él? Las respuestas finales para el resultado de cualquier medición deberían ser las mismas, pero durante los cálculos las diversas amplitudes y matrices usadas serán diferentes. ¿Cómo están relacionadas? Por ejemplo, si ambos comenzamos con el mismo ϕ , lo describiremos en función de las tres amplitudes $\langle iS | \phi \rangle$ para que ϕ vaya a nuestros estados base en la representación S , mientras que él lo describirá con las amplitudes $\langle jT | \phi \rangle$ para que el estado ϕ vaya a uno de los estados base en su representación T . ¿Cómo podemos comprobar que ambos estemos describiendo el mismo estado ϕ ? Podemos hacerlo con la regla general II de (5.27). Al remplazar a χ con cualquiera de sus estados jT , tenemos

$$\langle jT | \phi \rangle = \sum_i \langle jT | iS \rangle \langle iS | \phi \rangle. \quad (5.37)$$

Para relacionar las dos representaciones, sólo necesitamos dar los nueve números complejos de la matriz $\langle jT | iS \rangle$. Esta matriz se puede usar para convertir todas nuestras ecuaciones a la forma de nuestro compañero. Ésta nos dice cómo *transformar* de un conjunto de estados base a otro. (Por esta razón, $\langle jT | iS \rangle$ es llamada a veces “la matriz de transformación de la representación S a la representación T ”. ¡Palabras imponentes!)

En el caso de partículas con espín uno, para las cuales solamente tenemos tres estados base (para espines mayores, existen más), la situación matemática es análoga a lo que hemos visto en álgebra vectorial. Cada vector se puede representar mediante tres números —las componentes a lo largo de los ejes x , y y z —. Esto es, cada vector se puede descomponer en tres vectores “base”, que son vectores en la dirección de los tres ejes. Pero supongan que alguien decide usar un conjunto diferente de ejes — x' , y' y z' —. Estará usando números diferentes para representar a cualquier vector determinado. Sus cálculos parecerán diferentes, pero los resultados finales serán los mismos. Ya hemos considerado esta situación antes y conocemos las reglas para transformar vectores de un conjunto de ejes a otro.

Ustedes querrán comprobar cómo funcionan las transformaciones cuánticas intentando algunas; así que aquí les daremos, sin prueba, las matrices de transformación para convertir amplitudes para espín uno de una representación S a otra representación T , para diversas orientaciones relativas de los filtros S y T . (Les mostraremos en un capítulo posterior cómo derivar estos mismos resultados.)

Primer caso: El aparato T tiene el mismo eje y (a lo largo del cual se desplazan las partículas) que el aparato S , pero está girado alrededor del eje común y en un ángulo α (como en la figura 5-6). (Para ser más específicos, un conjunto de coordenadas x' , y' , z' se fija en el aparato T , relacionadas a las coordenadas x , y , z , del aparato S mediante: $z' = z \cos \alpha + x \sin \alpha$, $x' = x \cos \alpha - z \sin \alpha$, $y' = y$.) Entonces las amplitudes de la transformación son:

$$\begin{aligned} \langle +T | +S \rangle &= \frac{1}{2}(1 + \cos \alpha), \\ \langle 0T | +S \rangle &= -\frac{1}{\sqrt{2}} \sin \alpha, \\ \langle -T | +S \rangle &= \frac{1}{2}(1 - \cos \alpha), \\ \langle +T | 0S \rangle &= +\frac{1}{\sqrt{2}} \sin \alpha, \\ \langle 0T | 0S \rangle &= \cos \alpha, \\ \langle -T | 0S \rangle &= -\frac{1}{\sqrt{2}} \sin \alpha, \\ \langle +T | -S \rangle &= \frac{1}{2}(1 - \cos \alpha), \\ \langle 0T | -S \rangle &= +\frac{1}{\sqrt{2}} \sin \alpha, \\ \langle -T | -S \rangle &= \frac{1}{2}(1 + \cos \alpha). \end{aligned} \quad (5.38)$$

Segundo caso: El aparato T tiene el mismo eje z que S , pero está girado alrededor del eje z en un ángulo β . (La transformación de coordenadas es $z' = z$, $x' = x \cos \beta + y \sin \beta$, $y' = y \cos \beta - x \sin \beta$.) Las amplitudes para la transformación son entonces:

$$\begin{aligned}\langle +T | +S \rangle &= e^{+i\beta}, \\ \langle 0T | 0S \rangle &= 1, \\ \langle -T | -S \rangle &= e^{-i\beta}, \\ \text{las demás} &= 0.\end{aligned}\tag{5.39}$$

Noten que cualquier rotación de T se puede construir mediante las dos rotaciones descritas.

Si un estado ϕ es definido por los tres números

$$C_+ = \langle +S | \phi \rangle, \quad C_0 = \langle 0S | \phi \rangle, \quad C_- = \langle -S | \phi \rangle, \tag{5.40}$$

y el mismo estado es descrito desde el punto de vista de T por los tres números

$$C'_+ = \langle +T | \phi \rangle, \quad C'_0 = \langle 0T | \phi \rangle, \quad C'_- = \langle -T | \phi \rangle, \tag{5.41}$$

entonces los coeficientes $\langle jT | iS \rangle$ de (5.38) o de (5.39) nos dan la transformación que relaciona a C_i con C'_i . En otras palabras, las C_i son bastante parecidas a las componentes de un vector que parece distinto visto desde S o desde T .

Solamente para una partícula de espín uno —porque requiere *tres* amplitudes— la semejanza con un vector es muy estrecha. En cada caso, hay tres números que se deben transformar con cambios de coordenadas de una manera específica. De hecho, hay un conjunto de estados base *que se transforma exactamente como las tres componentes de un vector*. Las tres combinaciones

$$C_x = -\frac{1}{\sqrt{2}}(C_+ - C_-), \quad C_y = -\frac{i}{\sqrt{2}}(C_+ + C_-), \quad C_z = C_0 \tag{5.42}$$

transforman a C'_x , C'_y , y C'_z , del mismo modo que x , y , z se transforman en x' , y' , z' . [Pueden comprobar que esto es así usando las leyes de transformación (5.38) y (5.39).] Ahora ven por qué a una partícula con espín uno se le llama con frecuencia una “partícula vectorial”.

5-8 Otras situaciones

Comenzamos señalando que nuestra discusión sobre las partículas con espín uno serviría de prototipo para cualquier problema cuántico. La generalización solamente tiene que ver con el número de estados. En lugar de sólo tres estados base, cualquier situación particular puede incluir n estados base.† Nuestras leyes básicas en la ecuación (5.27) tienen exactamente la misma forma —entendiendo que i y j deben recorrer los n estados base—. Cualquier fenómeno se puede analizar con sólo dar las amplitudes para que comience en cada uno de los estados base y para que termine en cualquier otro de los estados base, y después sumar sobre el conjunto completo de estados base. Cualquier conjunto apropiado de estados base se puede usar, y si alguien desea utilizar un conjunto diferente, también está bien; los dos se pueden relacionar usando una matriz de transformación de $n \times n$. Más adelante insistiremos sobre tales transformaciones.

Finalmente, prometimos señalar qué hacer si los átomos llegan directamente de un horno, atraviesan algún aparato, digamos A , y luego se analizan con un filtro que selecciona al estado χ . Ustedes no saben en cuál estado ϕ se encuentran al comienzo. Quizá es mejor que no se preocupen por este problema ahora, y en su lugar se concentren en problemas que siempre comienzan con estados puros. Pero si insisten, aquí tienen la forma en que pueden tratar el problema.

† El número de estados base n puede ser, y generalmente lo es, infinito.

En primer lugar, necesitan ser capaces de hacer una conjetura razonable acerca de cómo están distribuidos los estados en los átomos que salen del horno. Por ejemplo, si no hubiese nada “especial” en el horno, podrían suponer de manera razonable que los átomos saldrían del horno con “orientaciones” aleatorias. Hablando de manera cuántica, esto corresponde a decir que no saben nada acerca de los estados, pero que un tercio de ellos está en el estado $(+S)$, un tercio está en el estado $(0S)$ y un tercio está en el estado $(-S)$. Para aquellos que se encuentran en el estado $(+S)$, la amplitud para atravesar es $\langle \chi | A | +S \rangle$ y la probabilidad es $|\langle \chi | A | +S \rangle|^2$, y lo mismo ocurre con los demás. La probabilidad total es entonces

$$\frac{1}{3}|\langle \chi | A | +S \rangle|^2 + \frac{1}{3}|\langle \chi | A | 0S \rangle|^2 + \frac{1}{3}|\langle \chi | A | -S \rangle|^2.$$

¿Por qué usamos S en vez de, por ejemplo, T ? El resultado es, sorprendentemente, el mismo sin importar qué escojamos para nuestra resolución inicial —siempre que estemos tratando con orientaciones completamente aleatorias—. Resulta del mismo modo que

$$\sum_i |\langle \chi | iS \rangle|^2 = \sum_j |\langle \chi | jT \rangle|^2$$

para cualquier χ . (Lo dejamos para que lo demuestren ustedes.)

Noten que *no* es correcto decir que el estado de entrada tiene las amplitudes $\sqrt{1/3}$ para estar en $(+S)$, $\sqrt{1/3}$ para estar en $(0S)$, y $\sqrt{1/3}$ para estar en $(-S)$; esto implicaría que podrían ocurrir ciertas interferencias. Simplemente no *saben* cuál es el estado inicial; tienen que pensar en términos de la probabilidad para que el sistema empiece en los diversos estados iniciales posibles, y entonces tomar un promedio ponderado sobre las diversas posibilidades.

Espín un medio[†]

6-1 Transformando amplitudes

En el capítulo anterior, utilizando un sistema de espín uno como ejemplo, delineamos los principios generales de la mecánica cuántica:

Cualquier estado ψ se puede describir en términos de un conjunto de estados base al dar las amplitudes para estar en cada uno de los estados base.

Las amplitudes para ir de cualquier estado a otro se pueden, en general, escribir como la suma de productos, siendo cada producto la amplitud para ir a uno de los estados base por la amplitud para ir desde ese estado base a la condición final, e incluyendo en la suma un término por cada estado base:

$$\langle \chi | \psi \rangle = \sum_i \langle \chi | i \rangle \langle i | \psi \rangle. \quad (6.1)$$

Los estados base son ortogonales —la amplitud para encontrarse en uno, si se está en otro, es cero—:

$$\langle i | j \rangle = \delta_{ij}. \quad (6.2)$$

La amplitud para ir de un estado a otro directamente es el complejo conjugado para ir a la inversa:

$$\langle \chi | \psi \rangle^* = \langle \psi | \chi \rangle. \quad (6.3)$$

También discutimos un poco acerca de que puede haber más de una base para los estados, y de que podemos usar la ecuación (6.1) para pasar de una base a otra. Supongamos, por ejemplo, que tenemos las amplitudes $\langle iS | \psi \rangle$ para encontrar al estado ψ en cada uno de los estados base i de un sistema de base S , pero luego decidimos que sería preferible describir el estado en términos de otro conjunto de estados base, digamos los estados j que pertenecen a la base T . En la fórmula general, ecuación (6.1), podemos sustituir jT por χ y obtenemos esta fórmula:

$$\langle jT | \psi \rangle = \sum_i \langle jT | iS \rangle \langle iS | \psi \rangle. \quad (6.4)$$

Las amplitudes para que el estado (ψ) esté en los estados base (jT) están relacionadas con las amplitudes para que esté en los estados base (iS) por medio del conjunto de coeficientes $\langle jT | iS \rangle$. Si hay N estados base, hay N^2 coeficientes. Ese conjunto de coeficientes a menudo se denomina “*matriz de transformación* para ir de la *representación S* a la *representación T*”. Esto parece matemáticamente difícil, pero, al cambiar un poco los nombres, podemos ver que realmente no es tan complicado. Si llamamos C_i a la amplitud para que el estado ψ esté en el estado base iS —esto es, $C_i = \langle iS | \psi \rangle$ — y llamamos C'_j a las amplitudes correspondientes al sistema de base T —esto es, $C'_j = \langle jT | \psi \rangle$ — entonces, la ecuación (6.4) se puede escribir en la siguiente forma

$$C'_j = \sum_i R_{ji} C_i, \quad (6.5)$$

6-1 Transformando amplitudes

6-2 Transformando a un sistema de coordenadas rotado

6-3 Rotaciones alrededor del eje z

6-4 Rotaciones de 180° y de 90° alrededor de y

6-5 Rotaciones alrededor de x

6-6 Rotaciones arbitrarias

[†] Este capítulo es una larga y abstracta digresión, y no introduce ninguna idea que no vayamos a obtener usando otros medios en capítulos posteriores. Pueden, por tanto, omitirlo y volver a él más tarde si están interesados.

donde R_{ji} significa lo mismo que $\langle jT|iS\rangle$. Cada amplitud C'_j es igual a una suma sobre toda i de uno de los coeficientes R_{ji} por cada amplitud C_i . Tiene la misma forma que la transformación de un vector de un sistema de coordenadas a otro.

Para no ser muy abstractos por mucho tiempo, les hemos dado algunos ejemplos de estos coeficientes para el caso del espín uno, y así pueden ver cómo usarlos en la práctica. Por otra parte, hay una cosa muy bella en la mecánica cuántica —por el mero hecho de que hay tres estados y por las propiedades de simetría del espacio ante rotaciones—, estos coeficientes se pueden obtener simplemente por medio de razonamiento abstracto. Mostrarles tales argumentos tan prematuramente tiene la desventaja de que se sumergirán en otro cúmulo de abstracciones antes de poner los pies “sobre la tierra”. No obstante, los argumentos son tan bellos, que lo vamos a hacer de todas maneras.

Les enseñaremos, en este capítulo, cómo se pueden obtener los coeficientes de transformación para partículas de espín un medio. Elegimos este caso, en vez del de espín uno, porque es algo más fácil. Nuestro problema es determinar los coeficientes R_{ji} para una partícula —un sistema atómico— que se separa en dos haces en un aparato de Stern-Gerlach. Vamos a deducir todos los coeficientes para la transformación de una representación a otra por puro razonamiento —más unas pocas suposiciones—. ¡*Algunas* suposiciones siempre son necesarias al usar razonamiento “puro”! Aunque los argumentos serán abstractos y un poco elaborados, el resultado al que llegaremos será relativamente simple de enunciar y fácil de comprender —y el resultado es lo más importante—. Pueden, si lo desean, considerarlo como una excursión cultural. Hemos, en realidad, arreglado que todos los resultados esenciales obtenidos aquí, se obtengan también de otra manera cuando los necesitemos en capítulos posteriores. Por ello, no teman perder el hilo de nuestro estudio de la mecánica cuántica si omiten por completo este capítulo, o lo estudian después. La excursión es “cultural” en el sentido de que intenta mostrar que los principios de la mecánica cuántica no sólo son interesantes, sino que son tan profundos que, al agregar pocas hipótesis acerca de la estructura del espacio, podemos inferir una gran cantidad de propiedades de los sistemas físicos. Además, es importante que sepamos de dónde provienen las diversas consecuencias de la mecánica cuántica, porque mientras nuestras leyes de la física estén incompletas —y sabemos que lo están— es interesante descubrir si las partes donde nuestras teorías dejan de concordar con el experimento son donde nuestro razonamiento es el mejor o donde es el peor. Hasta ahora, parece que donde nuestro razonamiento es más abstracto siempre se obtienen resultados correctos —concuerdan con el experimento—. Sólo cuando tratamos de hacer modelos específicos sobre la maquinaria interna de las partículas fundamentales y sus interacciones no podemos encontrar una teoría que concuerda con el experimento. La teoría que vamos a describir concuerda con los experimentos en todos los casos en que ha sido puesta a prueba —para las partículas extrañas así como para los electrones, los protones y demás—.

Antes de seguir, un comentario sobre un punto molesto pero interesante: no es posible determinar los coeficientes R_{ji} de manera única, porque siempre hay algo de arbitrariedad en las amplitudes de probabilidad. Si tienen un conjunto de amplitudes de cualquier tipo, digamos las amplitudes para llegar a algún lugar por todas las rutas posibles, y multiplican cada amplitud por el mismo factor de fase —digamos por $e^{i\delta}$ — obtienen otro conjunto que es tan adecuado como el anterior. Por consiguiente, siempre es posible hacer un cambio arbitrario en la fase de todas las amplitudes en un problema determinado, si así lo desean.

Supongan que calculamos una probabilidad escribiendo la suma de varias amplitudes, digamos $(A + B + C + \dots)$ y calculamos el cuadrado de la norma. Luego alguien más calcula lo mismo utilizando la suma de amplitudes $(A' + B' + C' + \dots)$ y calcula el cuadrado de la norma. Si todas las A' , B' , C' , etcétera, son iguales a las A , B , C , etcétera, excepto por un factor de fase $e^{i\delta}$, todas las probabilidades obtenidas tomando el cuadrado de la norma serán exactamente iguales, puesto que $(A' + B' + C' + \dots)$ es igual a $e^{i\delta}(A + B + C + \dots)$. O supongamos, por ejemplo, que estábamos calculando algo con la ecuación (6.1),

y que luego cambiamos repentinamente todas las fases de un cierto sistema de base. Cada amplitud $\langle i | \psi \rangle$ queda multiplicada por el mismo factor $e^{i\delta}$. Análogamente, las amplitudes $\langle i | \chi \rangle$ también cambiarían por $e^{i\delta}$, pero las amplitudes $\langle \chi | i \rangle$ son los complejos conjugados de las amplitudes $\langle i | \chi \rangle$; por lo tanto, éstas cambian por un factor $e^{-i\delta}$. Los más y menos $i\delta$ en los exponentes se cancelan y tenemos la misma expresión que teníamos antes. Luego, es una regla general que si cambiamos todas las amplitudes respecto a un sistema de base determinado por la misma fase —o incluso si cambiamos *todas* las amplitudes de un problema cualquiera por la misma fase— llegamos a los mismos resultados. Hay, por lo tanto, cierta libertad para escoger las fases de nuestra matriz de transformación. De vez en cuando haremos una elección arbitraria —por lo general siguiendo las convenciones que estén en uso—.

6-2 Transformando a un sistema de coordenadas rotado

Consideremos nuevamente el aparato de Stern-Gerlach “mejorado” descrito en el capítulo anterior. Un haz de partículas con espín un medio, que entra por la izquierda, se separa, en general, en *dos* haces, como se muestra en la figura 6-1. (Había tres haces para el caso de espín *uno*.) Como antes, los haces se juntan de nuevo, a no ser que uno o el otro sea bloqueado con un “tope” que intercepte al haz a mitad de camino. En la figura mostramos una flecha que apunta en la dirección en la que aumenta la *magnitud* del campo —digamos que hacia el polo del imán que tiene bordes agudos—. Tomamos esta flecha para representar *el eje “hacia arriba”* de cualquier aparato determinado. Está fijo respecto al aparato y nos permitirá indicar las orientaciones relativas cuando usemos varios aparatos juntos. También suponemos que la dirección del campo magnético de cada imán es siempre la misma respecto a la flecha.

Diremos que aquellos átomos que van al haz “superior” se encuentran en el estado (+) *respecto a ese aparato*, y los que van al haz “inferior” están en el estado (−). (No hay estado “cero” para las partículas con espín un medio.)

Ahora supongamos que ponemos dos de nuestros aparatos de Stern-Gerlach modificados en sucesión, como se muestra en la figura 6-2(a). El primero, al que llamamos *S*, se puede usar para preparar un estado (+*S*) puro o un estado (−*S*) puro al bloquear un haz o el otro. [Como se muestra en la figura, se prepara un estado (+*S*) puro.] Para cada condición, hay una cierta amplitud para que una partícula que sale de *S* esté en el haz (+*T*) o (−*T*) del segundo aparato. En realidad, sólo hay cuatro amplitudes: la amplitud para ir de (+*S*) a (+*T*), de (+*S*) a (−*T*), de (−*S*) a (+*T*) y de (−*S*) a (−*T*). Estas amplitudes son los cuatro coeficientes de la matriz de transformación R_{ji} para pasar de la representación *S* a la representación *T*. Podemos considerar que el primer aparato “prepara” un estado determinado en una representación, y que el segundo aparato “analiza” ese estado en términos de la segunda representación. El tipo de pregunta que buscamos responder, entonces, es ésta: si un átomo se ha preparado en un estado determinado —el (+*S*) por ejemplo— al bloquear uno de los haces en el aparato *S*, cuál es la probabilidad para que atraviese el segundo aparato *T*, si éste se ajusta para, digamos, el estado (−*T*). Por supuesto, el resultado dependerá de los ángulos entre los dos sistemas *S* y *T*.

Deberíamos explicar por qué tenemos la esperanza de encontrar los coeficientes R_{ji} por deducción. Ustedes saben que es casi imposible pensar que si una partícula tiene su espín alineado en la dirección +*z*, pueda haber alguna posibilidad de encontrar a la misma partícula con su espín apuntando en la dirección +*x* —o en cualquier otra dirección—. De hecho, *es* casi imposible, pero no del todo. Es tan imposible que *sólo hay una manera* de hacerlo, y ésa es la razón por la que podemos descubrir cuál es esa única manera.

El primer tipo de argumento que podemos dar es éste. Supongamos que tenemos un arreglo como el de la figura 6-2(a), en el cual tenemos dos aparatos *S* y *T*, con *T* inclinado un ángulo α respecto a *S*, y sólo dejamos que el haz (+) atraviese a *S*, y que el haz (−) atraviese a *T*. Deberíamos observar un cierto valor de la probabilidad para que las partículas que salen de *S* atraviesen a *T*.

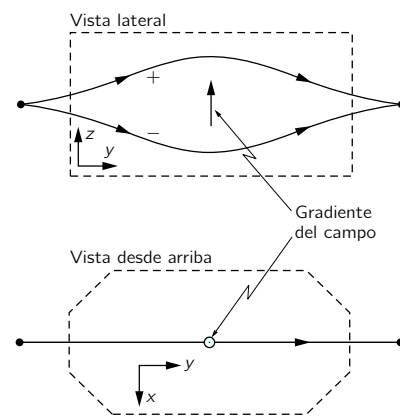


Figura 6-1. Vistas lateral y desde arriba de un aparato de Stern-Gerlach “mejorado” con los haces de una partícula con espín un medio.

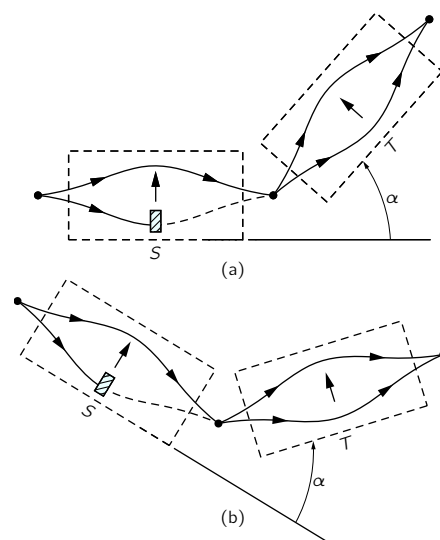


Figura 6-2. Dos experimentos equivalentes.

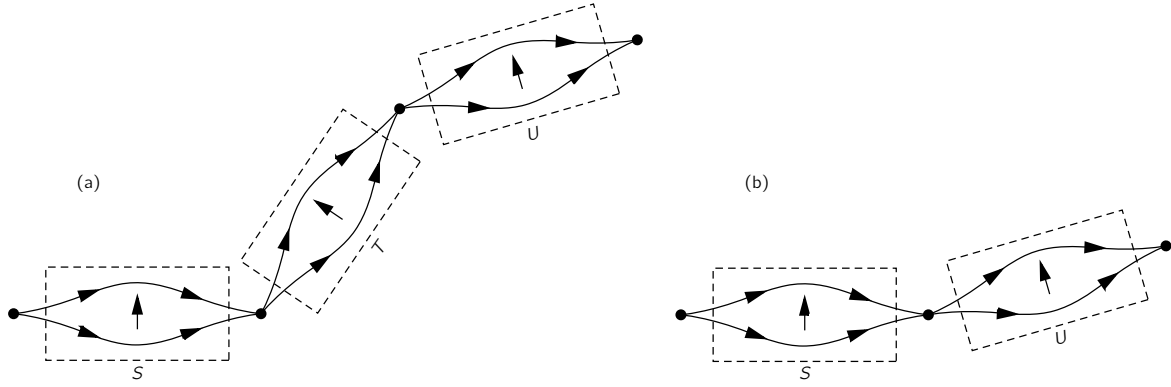


Figura 6-3. Si T está “abierto por completo”, (b) es equivalente a (a).

Ahora, supongamos que hacemos otra medición con el aparato de la figura 6-2(b). La orientación *relativa* de S y T es la misma, pero todo el sistema está en un ángulo diferente en el espacio. Queremos *suponer* que estos dos experimentos dan el mismo valor para la probabilidad para que una partícula en un estado puro respecto a S pase a algún estado determinado respecto a T . Estamos suponiendo, en otras palabras, que el resultado de cualquier experimento de este tipo es el mismo —que la *física* es la misma— sin importar cómo se encuentra orientado *todo* el aparato en el espacio. (Ustedes dirán, “Eso es obvio”. Pero *es* una suposición, y es “correcta” sólo si es lo que realmente ocurre.) Eso significa que los coeficientes R_{ji} sólo dependen de la relación en el espacio entre S y T , y no de la posición absoluta de S y T . Dicho de otro modo, R_{ji} sólo depende de la *rotación* que lleva a S hacia T , porque, evidentemente, lo que es igual en las figuras 6-2(a) y (b) es la rotación tridimensional que llevaría el aparato S a la misma orientación que el aparato T . Cuando la matriz de transformación R_{ji} sólo depende de una rotación, como lo hace aquí, se le denomina *matriz de rotación*.

Para nuestro siguiente paso necesitaremos una pieza más de información. Supongamos que agregamos un tercer aparato, al que podemos llamar U , que sigue de T formando un ángulo arbitrario, como en la figura 6-3(a). (Está empezando a verse horrible, pero eso es lo divertido del pensamiento abstracto —¡podemos hacer los experimentos más raros tan sólo dibujando líneas!—.) Ahora bien, ¿cuál es la transformación $S \rightarrow T \rightarrow U$? Lo que realmente queremos preguntar es la amplitud para ir de un estado respecto a S a algún otro estado respecto a U , cuando conocemos la transformación de S a T y de T a U . Estamos preguntando, entonces, sobre un experimento en el que ambos canales de T están abiertos. Podemos obtener la respuesta aplicando la ecuación (6.5) dos veces seguidas. Para ir de la representación S a la representación T tenemos

$$C'_j = \sum_i R_{ji}^{TS} C_i, \quad (6.6)$$

donde ponemos los superíndices TS en la R , para que podamos distinguirlas de los coeficientes R^{UT} que tendremos para ir de T a U .

Si suponemos que las amplitudes están en los estados base de la representación U , C''_k , podemos relacionarlas con las amplitudes en T usando la ecuación (6.5) una vez más, obtenemos

$$C''_k = \sum_j R_{kj}^{UT} C'_j. \quad (6.7)$$

Ahora podemos combinar las ecuaciones (6.6) y (6.7) para obtener la transformación a U directamente de S . Sustituyendo C'_j de la ecuación (6.6) en la ecuación (6.7), tenemos

$$C''_k = \sum_j R_{kj}^{UT} \sum_i R_{ji}^{TS} C_i. \quad (6.8)$$

O, como i no aparece en R_{kj}^{UT} , también podemos poner la suma sobre i al frente, y escribir

$$C_k'' = \sum_i \sum_j R_{kj}^{UT} R_{ji}^{TS} C_i. \quad (6.9)$$

Ésta es la fórmula para una transformación doble.

Noten, sin embargo, que mientras todos los haces en T no estén bloqueados, el estado que sale de T es el mismo que entró. Del mismo modo, podríamos haber hecho una transformación de la representación S de manera directa a la representación U . Sería lo mismo que poner el aparato U justo después de S , como en la figura 6-3(b). En ese caso, tendríamos que haber escrito

$$C_k'' = \sum_i R_{ki}^{US} C_i, \quad (6.10)$$

con los coeficientes R_{ki}^{US} que pertenecen a esta transformación. Ahora bien, está claro que las ecuaciones (6.9) y (6.10) deben dar las mismas amplitudes C_k'' , y que esto debe ser cierto sin importar cuál sea el estado original ϕ que nos dio las amplitudes C_i . Por lo que debe cumplirse

$$R_{ki}^{US} = \sum_j R_{kj}^{UT} R_{ji}^{TS}. \quad (6.11)$$

En otras palabras, para cualquier rotación $S \rightarrow U$ de una base de referencia, que se considera compuesta de dos rotaciones sucesivas $S \rightarrow T$ y $T \rightarrow U$, la matriz de rotación R_{ki}^{US} se puede obtener a partir de las matrices de las dos rotaciones parciales por medio de la ecuación (6.11). Si lo desean, pueden llegar a la ecuación (6.11) directamente de la ecuación (6.1), ya que sólo es una notación diferente para $\langle kU | iS \rangle = \sum_j \langle kU | jT \rangle \langle jT | iS \rangle$.

Para ser precisos, debemos agregar entre paréntesis los siguientes comentarios. No son tremendamente importantes, por ello, pueden pasar a la próxima sección si quieren. Lo que hemos dicho no es completamente correcto. Realmente no podemos decir que la ecuación (6.9) y la ecuación (6.10) deben dar *exactamente* las mismas amplitudes. Sólo la *física* debe ser la misma; todas las amplitudes podrían diferir en algún factor de fase común, como $e^{i\delta}$, sin cambiar el resultado de ningún cálculo sobre el mundo real. Por lo tanto, en vez de la ecuación (6.11), todo lo que podemos decir es que

$$e^{i\delta} R_{ki}^{US} = \sum_j R_{kj}^{UT} R_{ji}^{TS}, \quad (6.12)$$

donde δ es *alguna* constante real. Lo que este factor adicional $e^{i\delta}$ significa, por supuesto, es que las amplitudes que obtengamos si usamos la matriz R^{US} podrían diferir por el mismo factor de fase ($e^{-i\delta}$), de las amplitudes que obtendríamos empleando las dos rotaciones R^{UT} y R^{TS} . Sabemos que no importa si todas las amplitudes se cambian por la misma fase, por lo que simplemente podríamos ignorar este factor de fase si así lo quisiéramos. Resulta, sin embargo, que si definimos todas nuestras matrices de rotación de una manera determinada, este factor de fase adicional nunca aparecerá —la δ en la ecuación (6.12) siempre será cero—. Aunque no es importante para el resto de nuestra discusión, haremos una demostración rápida utilizando un teorema matemático sobre los determinantes. [Si aún no saben mucho sobre determinantes, no se preocupen por la demostración y pasen simplemente a la definición (6.15).]

En primer lugar, tenemos que decir que la ecuación (6.11) es la definición matemática del “producto” de dos matrices. (Sólo que es conveniente poder decir: “ R^{US} es el producto de R^{UT} y R^{TS} ”). En segundo lugar, hay un teorema en matemáticas —que pueden probar fácilmente para las matrices de dos por dos que tenemos aquí— que dice que el determinante de un “producto” de dos matrices

es el producto de sus determinantes. Aplicando este teorema a la ecuación (6.12) obtenemos

$$e^{i2\delta} (\text{Det } R^{US}) = (\text{Det } R^{UT}) \cdot (\text{Det } R^{TS}). \quad (6.13)$$

(Quitamos los subíndices, porque no nos dicen nada útil.) Sí, el 2δ es correcto. Recuerden que estamos tratando con matrices de dos por dos; cada término de la matriz R_{ki}^{US} se multiplica por $e^{i\delta}$, por lo que cada producto en el determinante —que tiene *dos* factores— queda multiplicado por $e^{i2\delta}$. Ahora dividamos (6.12) entre la raíz cuadrada de la ecuación (6.13); obtenemos

$$\frac{R_{ki}^{US}}{\sqrt{\text{Det } R^{US}}} = \sum_j \frac{R_{kj}^{UT}}{\sqrt{\text{Det } R^{UT}}} \frac{R_{ji}^{TS}}{\sqrt{\text{Det } R^{TS}}}. \quad (6.14)$$

El factor de fase adicional ha desaparecido.

Ahora, ocurre que si queremos que todas nuestras amplitudes en cualquier representación estén normalizadas (lo cual significa que $\sum_i \langle \phi | i \rangle \langle i | \phi \rangle = 1$), todas las matrices de rotación tendrán determinantes que son exponenciales imaginarios puros, como $e^{i\alpha}$. (No lo demostraremos; verán que siempre ocurre así.) Entonces podemos, si queremos, hacer que todas nuestras matrices de rotación R tengan una fase única haciendo que $\text{Det } R = 1$. Esto se hace así. Supongamos que hemos encontrado una matriz de rotación R de algún modo arbitrario. Tomamos como regla “convertirla” a la “forma estándar” definiendo

$$R_{\text{estándar}} = \frac{R}{\sqrt{\text{Det } R}}. \quad (6.15)$$

Podemos hacerlo porque sólo estamos multiplicando cada término de R por el mismo factor de fase, para obtener las fases que queremos. En lo siguiente, supondremos siempre que hemos llevado todas las matrices a la “forma estándar”; entonces podemos usar la ecuación (6.11) sin tener ningún factor de fase adicional.

6-3 Rotaciones alrededor del eje z

Ahora ya estamos listos para encontrar la matriz de transformación R_{ji} entre dos representaciones diferentes. Con nuestra regla para componer rotaciones y nuestra suposición de que el espacio no tiene ninguna dirección privilegiada, tenemos las claves que necesitamos para encontrar la matriz para cualquier rotación arbitraria. Hay sólo *una* solución. Comencemos con la transformación correspondiente a una rotación alrededor del eje z . Supongamos que tenemos dos aparatos, S y T , colocados en serie en línea recta con sus ejes paralelos apuntando hacia afuera de la página, como se muestra en la figura 6-4(a). Tomamos nuestro “eje z ” en esta dirección. Es seguro que si el haz va hacia “arriba” (hacia $+z$) en el aparato S , hará lo mismo en el aparato T . Análogamente, si va hacia abajo en S , también irá hacia abajo en T . Sin embargo, supongamos que el aparato T está ubicado en otro ángulo, pero aún con sus ejes paralelos a los ejes de S ,

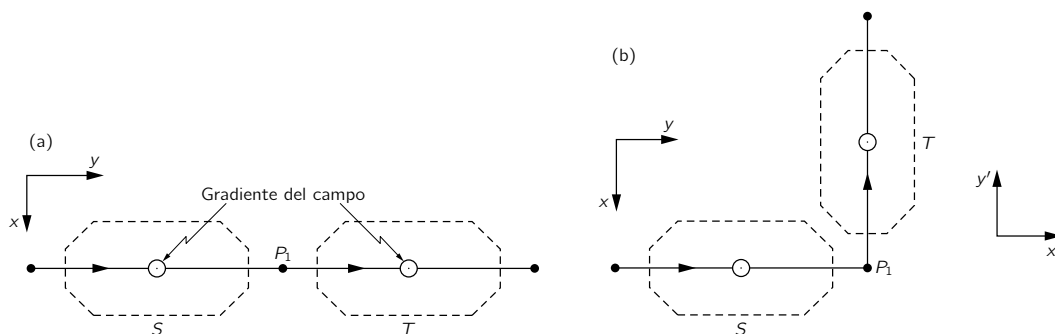


Figura 6-4. Rotación de 90° alrededor del eje z .

como en la figura 6-4(b). De manera intuitiva, dirían que a un haz (+) en S corresponde todavía a un haz (+) en T , porque los campos y los gradientes de los campos están aún en la misma dirección física. Y eso sería bastante correcto. También, a un haz (−) en S le seguiría correspondiendo un haz (−) en T . El mismo resultado se aplica para cualquier orientación de T en el plano xy de S . ¿Qué nos dice esto sobre la relación entre $C'_+ = \langle +T | \psi \rangle$, $C'_- = \langle -T | \psi \rangle$ y $C_+ = \langle +S | \psi \rangle$, $C_- = \langle -S | \psi \rangle$? Ustedes podrían concluir que cualquier rotación alrededor del eje z del “marco de referencia” para los estados base no cambia las amplitudes para estar “hacia arriba” y “hacia abajo”. Podríamos escribir $C'_+ = C_+$ y $C'_- = C_-$ —pero esto es *incorrecto*—. Lo único que *podemos* concluir es que para dichas rotaciones las probabilidades para estar en el haz “hacia arriba” son iguales en los aparatos S y T . Esto es,

$$|C'_+| = |C_+| \quad \text{y} \quad |C'_-| = |C_-|.$$

No podemos decir que las *fases* de las amplitudes referidas al aparato T no puedan ser distintas para las dos orientaciones diferentes mostradas en (a) y (b) de la figura 6-4.

Los dos aparatos en (a) y en (b) de la figura 6-4 son, de hecho, diferentes, como se puede ver del modo siguiente. Supongamos que ponemos un aparato frente a S que produce un estado $(+x)$ puro. (El eje x apunta hacia la parte inferior de la figura.) Dichas partículas se separarán en los haces $(+z)$ y $(-z)$ en S , pero los dos haces se recombinarán para dar de nuevo un estado $(+x)$ en P_1 —la salida de S —. Lo mismo ocurre nuevamente en T . Si seguimos a T con un tercer aparato U , cuyo eje esté en la dirección $(+x)$ y, como se muestra en la figura 6-5(a), todas las partículas van al haz (+) de U . Imaginemos ahora lo que ocurre si giramos T y U *conjuntamente* 90° hasta la posición mostrada en la figura 6-5(b). También aquí el aparato T deja salir exactamente lo que recibe, por lo que las partículas que entran en U están en un estado $(+x)$ respecto a S . Pero ahora U las analiza para el estado $(+y)$ respecto a S , lo cual es diferente. (Por simetría, esperamos que sólo la mitad de las partículas logre pasar.)

¿Qué pudo haber cambiado? Los aparatos T y U aún tienen la misma relación *física* entre sí. ¿Puede cambiar la *situación física* sólo porque T y U tienen una orientación diferente? Nuestra suposición original es que no tiene que cambiar. Debe ser que las *amplitudes* respecto a T son diferentes en los dos casos mostrados en la figura 6-5 —y, por lo tanto, también en la figura 6-4—. Debe haber alguna manera para que la partícula sepa que ha doblado la esquina en P_1 . ¿Cómo podría saberlo? Bueno, todo lo que hemos decidido es que las *magnitudes* de C'_+ y C_+ son iguales en ambos casos, pero podrían —en realidad, *deben*— tener *fases* diferentes. Concluimos que C'_+ y C_+ deben estar relacionados mediante

$$C'_+ = e^{i\lambda} C_+,$$

y que C'_- y C_- están relacionados mediante

$$C'_- = e^{i\mu} C_-,$$

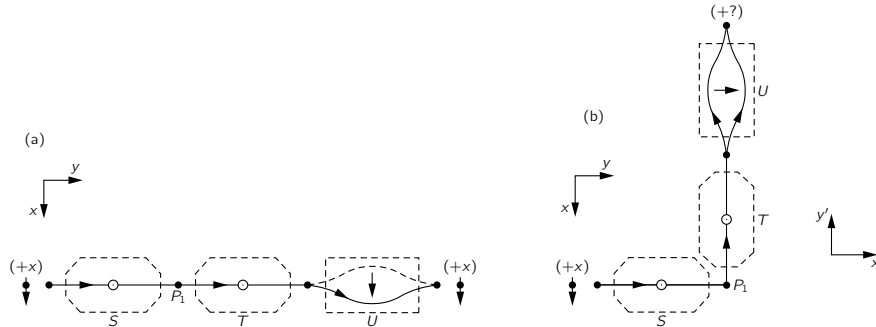


Figura 6-5. La partícula que está en estado $(+x)$ se comporta de un modo diferente en (a) y en (b).

donde λ y μ son números reales que tienen, de alguna manera, que estar relacionados con el ángulo entre S y T .

Lo único que podemos decir por el momento sobre λ y μ es que no deben ser iguales [excepto en el caso especial mostrado en la figura 6-5(a), cuando T tiene la misma orientación que S]. Hemos visto que cambios de fase iguales para todas las amplitudes no tienen ninguna consecuencia física. Por la misma razón, podemos sumar la misma cantidad arbitraria a λ y a μ sin cambiar nada. Por lo tanto, tenemos permitido *escoger* que λ y μ sean iguales a más y menos el mismo número. Es decir, siempre podemos hacer

$$\lambda' = \lambda - \frac{(\lambda + \mu)}{2}, \quad \mu' = \mu - \frac{(\lambda + \mu)}{2}.$$

Entonces

$$\lambda' = \frac{\lambda}{2} - \frac{\mu}{2} = -\mu'.$$

Así que adoptamos la convención† de que $\mu = -\lambda$. Tenemos entonces la regla general de que, para una rotación del aparato de referencia en algún ángulo alrededor del eje z , la transformación es

$$C'_+ = e^{+i\lambda} C_+, \quad C'_- = e^{-i\lambda} C_-. \quad (6.16)$$

Las normas son iguales, sólo las fases son diferentes. Estos factores de fase son responsables de los resultados diferentes en los dos experimentos de la figura 6-5.

Ahora queremos conocer la ley que relaciona a λ con el ángulo entre S y T . Ya conocemos la respuesta en un caso. Si el ángulo es cero, λ es cero. Ahora *suponemos* que el cambio de fase λ es una función continua del ángulo ϕ que hay entre S y T (véase la figura 6-4) cuando ϕ tiende a cero —lo cual parece razonable—. En otras palabras, si rotamos a T respecto a la recta que pasa por S en un ángulo pequeño ϵ , también λ es una cantidad pequeña, digamos $m\epsilon$, donde m es algún número. Lo escribimos así porque se puede demostrar que λ debe ser proporcional a ϵ . Supongamos que después de T ponemos otro aparato T' , el cual forma un ángulo ϵ con T , y, por lo tanto, un ángulo 2ϵ con S . Entonces, respecto a T , tenemos

$$C'_+ = e^{i\lambda} C_+,$$

y respecto a T' tenemos

$$C''_+ = e^{i\lambda} C'_+ = e^{i2\lambda} C_+.$$

Pero sabemos que obtendríamos el mismo resultado si pusiéramos a T' inmediatamente después de S . Luego entonces, cuando se duplica el ángulo, la fase se duplica. Es evidente que podemos extender este argumento y armar cualquier rotación con una sucesión de rotaciones infinitesimales. Concluimos que, para *cualquier* ángulo ϕ , λ es proporcional al ángulo. En consecuencia, podemos escribir $\lambda = m\phi$.

El resultado general que se obtiene es, entonces, que rotando T alrededor del eje z en un ángulo ϕ respecto a S

$$C'_+ = e^{im\phi} C_+, \quad C'_- = e^{-im\phi} C_-. \quad (6.17)$$

Para el ángulo ϕ , y para todas las rotaciones de las que hablemos en adelante, adoptamos la convención de que una rotación *positiva* es una rotación *a la derecha* alrededor de la dirección positiva del eje de referencia. Un ϕ positivo tiene el sentido de rotación de un tornillo derecho que avanza en la dirección z positiva.

Ahora tenemos que encontrar cuánto vale m . Primero, podríamos intentar este argumento: supongamos que se rota T en 360° ; es claro que, entonces, está de vuelta en cero grados, y debemos tener $C'_+ = C_+$ y $C'_- = C_-$, o lo que es lo mismo, $e^{im2\pi} = 1$. Obtenemos que $m = 1$. *¡Este argumento es incorrecto!* Para verlo, consideremos que se rota T en 180° . Si m fuera igual a 1, tendríamos

† Diciéndolo de otro modo, sólo estamos poniendo la transformación en la “forma estándar” mencionada en la sección 6-2, utilizando la ecuación (6.15).

$C'_+ = e^{i\pi}C_+ = -C_+$ y $C'_- = e^{-i\pi}C_- = -C_-$. Sin embargo, éste es justo el estado *original* de nuevo. Ambas amplitudes sólo están multiplicadas por -1 , lo cual da el sistema físico original. (De nuevo se trata de un cambio de fase común.) Esto significa que, si el ángulo entre T y S en la figura 6-5(b) se aumenta a 180° , el sistema (respecto a T) sería indistinguible de la situación con cero grados, y las partículas irían nuevamente al estado $(+)$ del aparato U . En 180° , no obstante, el estado $(+)$ del aparato U es el estado $(-x)$ del aparato S original. Por lo que un estado $(+x)$ se convertiría en un estado $(-x)$. Pero no hemos hecho nada para *cambiar* al estado original; la respuesta está mal. No podemos tener $m = 1$.

Se debe cumplir que una rotación de 360° , y de *ningún otro ángulo menor*, reproduce el mismo estado físico. Esto ocurre si $m = \frac{1}{2}$. Entonces, y sólo entonces, el primer ángulo que reproduce el mismo estado *físico* será $\phi = 360^\circ$.[†] Esto da

$$\left. \begin{aligned} C'_+ &= -C_+ \\ C'_- &= -C_- \end{aligned} \right\} 360^\circ \text{ alrededor del eje } z. \quad (6.18)$$

Es muy curioso decir que, si el aparato se gira 360° , se obtienen nuevas amplitudes. Realmente no son nuevas, sin embargo, porque el cambio común de signo no da ninguna situación física diferente. Si alguien decide cambiar todos los signos de las amplitudes por creer que ha girado 360° , no estaría mal; obtendría la misma situación física.[‡] Así que nuestra respuesta final es que, si conocemos las amplitudes C_+ y C_- para partículas con espín un medio respecto a un marco de referencia S , y luego usamos un sistema de base referido a T , el cual se obtiene de S por medio de una rotación de ϕ alrededor del eje z , las nuevas amplitudes están dadas en función de las antiguas por

$$\left. \begin{aligned} C'_+ &= e^{i\phi/2}C_+ \\ C'_- &= e^{-i\phi/2}C_- \end{aligned} \right\} \phi \text{ alrededor de } z. \quad (6.19)$$

6-4 Rotaciones de 180° y de 90° alrededor de y

A continuación, trataremos de encontrar la transformación para una rotación de T , respecto a S de 180° , alrededor de un eje *perpendicular* al eje z —por ejemplo, alrededor del eje y —. (Hemos definido los ejes de coordenadas en la figura 6-1.) En otras palabras, empezamos con dos equipos de Stern-Gerlach idénticos, con el segundo, T , “al revés” respecto al primero, S , como en la figura 6-6. Si ahora consideramos a nuestras partículas como pequeños dipolos magnéticos, una partícula que está en el estado $(+S)$ —de modo que va por el camino “de arriba” en el primer aparato—también tomará el camino “de arriba” en el segundo, por lo que estará en el estado *menos* respecto a T . (En el aparato invertido T , los gradientes y la dirección del campo se invierten; para una partícula cuyo momento magnético está en una dirección determinada, la fuerza no cambia.) De todos modos, lo que está “hacia arriba” respecto a S estará “hacia abajo” respecto a T . Para estas posiciones relativas de S y T sabemos, entonces, que la transformación debe dar

$$|C'_+| = |C_-|, \quad |C'_-| = |C_+|.$$

Como antes, no podemos descartar algunos factores de fase adicionales; podríamos tener (para una rotación de 180° alrededor del eje y)

$$C'_+ = e^{i\beta}C_- \quad \text{y} \quad C'_- = e^{i\gamma}C_+, \quad (6.20)$$

donde β y γ se deben determinar.

[†] Parece que $m = -\frac{1}{2}$ también funcionaría. Sin embargo, vemos en la ecuación (6.17) que el cambio de signo sólo redefine la notación para una partícula con espín hacia arriba.

[‡] Además, si se ha rotado algo por medio de una sucesión de rotaciones pequeñas, cuyo resultado total es regresar a la orientación original, es posible definir el concepto de que ha sido rotado 360° —como algo distinto de una rotación total cero—si han seguido el rastro de toda la historia. (Un dato interesante, esto *no* es cierto para una rotación total de 720° .)

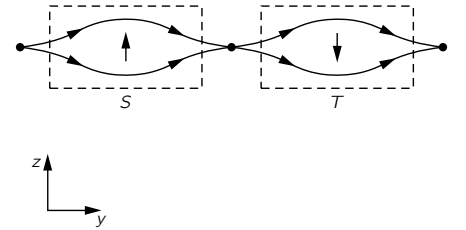


Figura 6-6. Una rotación de 180° alrededor del eje y .

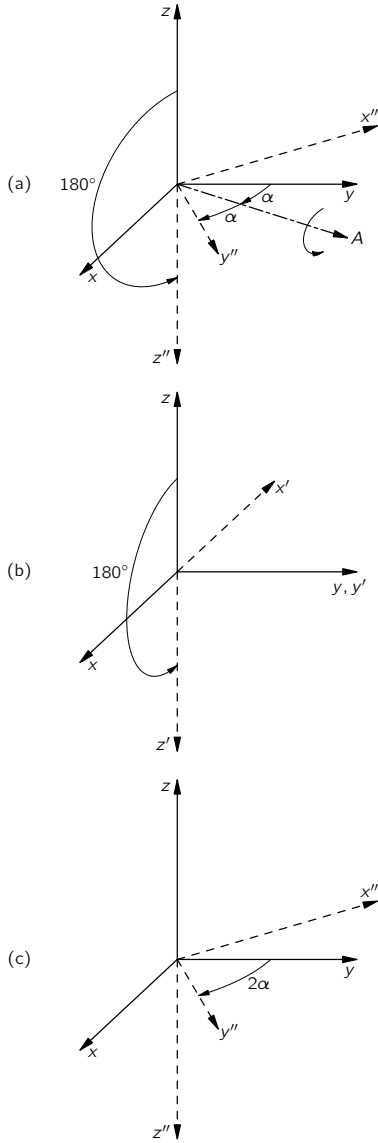


Figura 6-7. Una rotación de 180° alrededor del eje A es equivalente a una rotación de 180° alrededor de y , seguida de una rotación alrededor de z' .

¿Qué pasa con una rotación de 360° alrededor del eje y ? Ya sabemos la respuesta para una rotación de 360° alrededor del eje z —la amplitud para estar en cualquier estado cambia de signo—. Una rotación de 360° alrededor de cualquier eje siempre nos regresa a la posición original. Debe ser que para *cualquier* rotación de 360° , el resultado es el mismo que para una rotación de 360° alrededor del eje z —simplemente, todas las amplitudes cambian de signo—. Supongamos ahora que realizamos dos rotaciones sucesivas de 180° alrededor de y —utilizando la ecuación (6.20)— debemos obtener el resultado indicado en la ecuación (6.18). En otras palabras,

$$C''_+ = e^{i\beta} C'_+ = e^{i\beta} e^{i\gamma} C_+ = -C_+ \quad (6.21)$$

y

$$C''_- = e^{i\gamma} C'_- = e^{i\gamma} e^{i\beta} C_- = -C_-.$$

Esto significa que

$$e^{i\beta} e^{i\gamma} = -1 \quad \text{o} \quad e^{i\gamma} = -e^{-i\beta}.$$

Por lo que la transformación correspondiente a una rotación de 180° alrededor del eje y se puede escribir en la forma

$$C'_+ = e^{i\beta} C_-, \quad C'_- = -e^{-i\beta} C_+. \quad (6.22)$$

Los argumentos que acabamos de emplear se aplican también a rotaciones de 180° alrededor de *cualquier* eje en el plano xy , aunque por supuesto, distintos ejes pueden dar valores diferentes para β . Sin embargo, es la única manera en que pueden diferir. Ahora, hay una cierta arbitrariedad en el número β , pero una vez especificado para un eje de rotación en el plano xy , queda determinado para cualquier otro eje. Es *convención* elegir $\beta = 0$ para una rotación de 180° alrededor del eje y .

Para mostrar que podemos hacer esta elección, supongamos que β no fuera igual a cero para una rotación alrededor del eje y ; entonces podemos demostrar que hay algún *otro* eje en el plano xy , para el cual el factor de fase correspondiente *sí* es cero. Encontremos el factor de fase β_A para un eje A que forma el ángulo α con el eje y , como se muestra en la figura 6-7(a). (Por claridad, la figura está dibujada con α igual a un número negativo, pero eso no importa.) Si ahora tomamos un aparato T , que inicialmente está alineado con el aparato S , y luego lo rotamos 180° alrededor del eje A , sus ejes —que llamaremos x'' , y'' y z'' — estarán como se muestra en la figura 6-7(a). Las amplitudes respecto a T serán entonces

$$C''_+ = e^{i\beta_A} C_-, \quad C''_- = -e^{-i\beta_A} C_+. \quad (6.23)$$

Podemos pensar ahora en llegar a la misma orientación por medio de las dos rotaciones sucesivas mostradas en las partes (b) y (c) de la figura. Primero, imaginamos un aparato U que se rota 180° respecto a S alrededor del eje y . Los ejes x' , y' y z' de U quedan como se muestra en la figura 6-7(b), y las amplitudes respecto a U están dadas por la ecuación (6.22).

Vean ahora que podemos ir de U a T mediante una rotación alrededor del “eje z ” de U , a saber alrededor de z' , como se muestra en la figura 6-7(c). En la misma figura se puede ver que el ángulo necesario es dos veces el ángulo α , pero en sentido opuesto (respecto a z'). Usando la transformación (6.19) con $\phi = -2\alpha$, obtenemos

$$C''_+ = e^{-i\alpha} C'_+, \quad C''_- = e^{+i\alpha} C'_-. \quad (6.24)$$

Al combinar las ecuaciones (6.24) y (6.22), llegamos a

$$C''_+ = e^{i(\beta-\alpha)} C_-, \quad C''_- = -e^{-i(\beta-\alpha)} C_+. \quad (6.25)$$

Estas amplitudes deben ser, por supuesto, iguales a las de la ecuación (6.23). Por lo que β_A debe estar relacionado con α y β mediante

$$\beta_A = \beta - \alpha. \quad (6.26)$$

Esto significa que si el ángulo α entre el eje A y el eje y (de S) es igual a β , la transformación correspondiente a una rotación de 180° alrededor de A tendrá $\beta_A = 0$.

Entonces, como hay *algún* eje perpendicular al eje z que tiene $\beta = 0$, bien podemos tomarlo como el eje y . Es sólo cuestión de *convención*, y adoptamos la que es de uso general. Nuestro resultado es: para una rotación de 180° alrededor del eje y tenemos

$$\left. \begin{array}{l} C'_+ = C_- \\ C'_- = -C_+ \end{array} \right\} 180^\circ \text{ alrededor de } y. \quad (6.27)$$

Ya que estamos considerando al eje y , busquemos la matriz de transformación para una rotación de 90° alrededor de y . Podemos encontrarla, porque sabemos que dos rotaciones sucesivas de 90° alrededor del mismo eje equivalen a una rotación de 180° . Comenzamos escribiendo la transformación correspondiente a 90° en la forma más general:

$$C'_+ = aC_+ + bC_-, \quad C'_- = cC_+ + dC_-. \quad (6.28)$$

Una segunda rotación de 90° alrededor del mismo eje tendrá los mismos coeficientes:

$$C''_+ = aC'_+ + bC'_-, \quad C''_- = cC'_+ + dC'_-. \quad (6.29)$$

Combinamos las ecuaciones (6.28) y (6.29), obtenemos

$$\begin{aligned} C''_+ &= a(aC_+ + bC_-) + b(cC_+ + dC_-), \\ C''_- &= c(aC_+ + bC_-) + d(cC_+ + dC_-). \end{aligned} \quad (6.30)$$

Sin embargo, de la ecuación (6.27) sabemos que

$$C''_+ = C_-, \quad C''_- = -C_+,$$

por lo que debemos tener

$$\begin{aligned} ab + bd &= 1, \\ a^2 + bc &= 0, \\ ac + cd &= -1, \\ bc + d^2 &= 0. \end{aligned} \quad (6.31)$$

Estas cuatro ecuaciones son suficientes para determinar nuestras cuatro incógnitas a , b , c y d . No es difícil de hacer. Observemos las ecuaciones segunda y cuarta. Deducimos que $a^2 = d^2$, lo cual significa que $a = d$, o que $a = -d$. Pero $a = -d$ no sirve, porque entonces la primera ecuación no sería correcta. Entonces $d = a$. Usando este resultado, tenemos inmediatamente que $b = 1/2a$, y que $c = -1/2a$. Ahora tenemos todo en función de a . Escribiendo, por ejemplo, la segunda ecuación en términos de a , tenemos que

$$a^2 - \frac{1}{4a^2} = 0 \quad \text{o} \quad a^4 = \frac{1}{4}.$$

Esta ecuación tiene cuatro soluciones diferentes, pero sólo dos de ellas dan el valor estándar para el determinante. Podríamos muy bien elegir $a = 1/\sqrt{2}$; luego†

$$\begin{aligned} a &= 1/\sqrt{2}, & b &= 1/\sqrt{2}, \\ c &= -1/\sqrt{2}, & d &= 1/\sqrt{2}. \end{aligned}$$

En otras palabras, para dos aparatos S y T , con T rotado respecto a S en 90° alrededor del eje y , la transformación es

$$\left. \begin{array}{l} C'_+ = \frac{1}{\sqrt{2}}(C_+ + C_-) \\ C'_- = \frac{1}{\sqrt{2}}(-C_+ + C_-) \end{array} \right\} 90^\circ \text{ alrededor de } y. \quad (6.32)$$

† La otra solución cambia los signos de a , b , c y d y corresponde a una rotación de -270° .

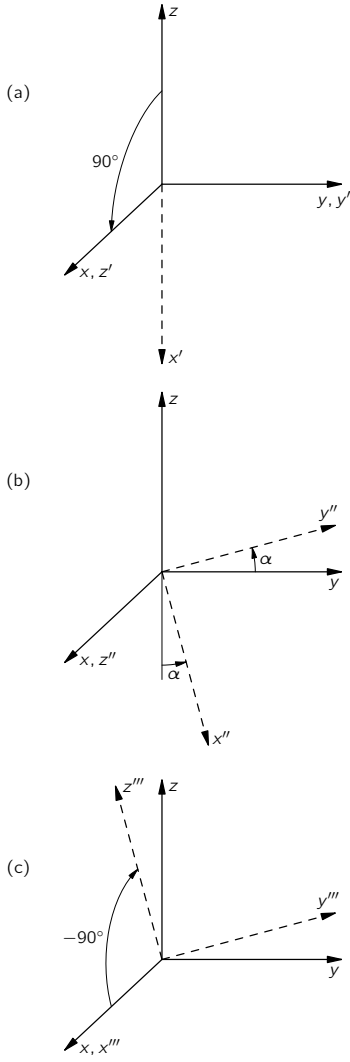


Figura 6-8. Una rotación en un ángulo α en el eje x es equivalente a: (a) una rotación de 90° alrededor de y , seguida por (b) una rotación de α alrededor de z' , seguida de (c) una rotación de -90° alrededor de y'' .

Podemos, por supuesto, despejar C_+ y C_- de estas ecuaciones, lo cual nos dará la transformación correspondiente a una rotación de *menos* 90° alrededor de y . Intercambiando las primas, concluimos que

$$\left. \begin{aligned} C'_+ &= \frac{1}{\sqrt{2}} (C_+ - C_-) \\ C'_- &= \frac{1}{\sqrt{2}} (C_+ + C_-) \end{aligned} \right\} -90^\circ \text{ alrededor de } y. \quad (6.33)$$

6-5 Rotaciones alrededor de x

Puede que estén pensando: “Esto se está poniendo ridículo. ¿Qué es lo que van a hacer después, 47° alrededor de y , luego 33° alrededor de x y así sucesivamente por siempre?” No, ya casi hemos terminado. Sólo con dos de las transformaciones que tenemos -90° alrededor de y , y un ángulo arbitrario alrededor de z (que es lo primero que hicimos primero si lo recuerdan)—podemos generar cualquier rotación.

Por ejemplo, supongamos que queremos el ángulo α alrededor de x . Sabemos tratar al ángulo α alrededor de z , pero ahora lo queremos alrededor de x . ¿Cómo lo lograremos? Primero, giramos el eje z hacia abajo hasta coincidir con x —lo cual es una rotación de $+90^\circ$ alrededor de y —, como se muestra en la figura 6-8. Después giramos un ángulo α alrededor del eje z' . Después rotamos -90° alrededor del nuevo eje y'' . El resultado final de estas tres rotaciones es el mismo que girar alrededor de x en un ángulo α . Es una propiedad del espacio.

(Estos hechos de la combinación de rotaciones, y lo que éstas producen, son difíciles de entender intuitivamente. Es bastante extraño, porque vivimos en tres dimensiones, pero nos resulta difícil ver lo que ocurre si damos vueltas de este modo y luego de aquel otro. Quizá, si fuésemos peces o aves y tuviéramos una percepción real de qué ocurre al dar volteretas en el espacio, nos daríamos cuenta mejor de dichas propiedades.)

De todas maneras, construyamos la transformación correspondiente a una rotación en un ángulo α alrededor del eje x utilizando lo que sabemos. Para la primera rotación de $+90^\circ$ alrededor de y , las amplitudes se transforman de acuerdo con la ecuación (6.32). Si llamamos x' , y' y z' a los ejes rotados, entonces la siguiente rotación en un ángulo α alrededor de z' nos lleva a un marco x'' , y'' , z'' , para los cuales

$$C''_+ = e^{i\alpha/2} C'_+, \quad C''_- = e^{-i\alpha/2} C'_-.$$

La última rotación de -90° alrededor de y'' nos lleva a x''' , y''' , z''' ; según (6.33),

$$C'''_+ = \frac{1}{\sqrt{2}} (C''_+ - C''_-), \quad C'''_- = \frac{1}{\sqrt{2}} (C''_+ + C''_-).$$

Al combinar estas dos transformaciones obtenemos

$$\begin{aligned} C'''_+ &= \frac{1}{\sqrt{2}} (e^{+i\alpha/2} C'_+ - e^{-i\alpha/2} C'_-), \\ C'''_- &= \frac{1}{\sqrt{2}} (e^{+i\alpha/2} C'_+ + e^{-i\alpha/2} C'_-). \end{aligned}$$

Con las ecuaciones (6.32) para C'_+ y C'_- , obtenemos la transformación completa:

$$\begin{aligned} C'''_+ &= \frac{1}{2} \{ e^{+i\alpha/2} (C_+ + C_-) - e^{-i\alpha/2} (-C_+ + C_-) \}, \\ C'''_- &= \frac{1}{2} \{ e^{+i\alpha/2} (C_+ + C_-) + e^{-i\alpha/2} (-C_+ + C_-) \}. \end{aligned}$$

Podemos escribir estas fórmulas en una manera más simple recordando que

$$e^{i\theta} + e^{-i\theta} = 2 \cos \theta, \quad \text{y} \quad e^{i\theta} - e^{-i\theta} = 2i \sin \theta.$$

Obtenemos

$$\left. \begin{aligned} C_+''' &= \left(\cos \frac{\alpha}{2} \right) C_+ + i \left(\sin \frac{\alpha}{2} \right) C_- \\ C_-''' &= i \left(\sin \frac{\alpha}{2} \right) C_+ + \left(\cos \frac{\alpha}{2} \right) C_- \end{aligned} \right\} \alpha \text{ alrededor de } x. \quad (6.34)$$

Aquí está nuestra transformación para una rotación alrededor del eje x en *cualquier* ángulo α . Sólo es un poco más complicada que las otras.

6-6 Rotaciones arbitrarias

Ahora podemos ver cómo tratar con *cualquier* ángulo. Primero, observen que cualquier orientación relativa de dos sistemas de coordenadas se puede describir con tres ángulos, como se muestra en la figura 6-9. Si tenemos un conjunto de ejes x' , y' y z' , orientados de cualquier manera respecto a x , y y z , podemos describir la relación entre los sistemas por medio de los tres ángulos de Euler, α , β y γ , que definen las tres rotaciones sucesivas que llevan al sistema x , y , z hacia el sistema x' , y' , z' . Comenzando en x , y , z , rotamos el sistema en un ángulo β alrededor del eje z , con esto el eje x va a una línea x_1 . Luego, rotamos en α grados alrededor de este eje x temporal, para llevar z hacia z' . Finalmente, una rotación alrededor del nuevo eje z (es decir, z') en un ángulo γ llevará el eje x hacia x' y el eje y a y' .[†] Conocemos las transformaciones correspondientes a cada una de las tres rotaciones —están dadas en las ecuaciones (6.19) y (6.34)—. Combinándolas en el orden apropiado, obtenemos

$$\begin{aligned} C_+' &= \cos \frac{\alpha}{2} e^{i(\beta+\gamma)/2} C_+ + i \sin \frac{\alpha}{2} e^{-i(\beta-\gamma)/2} C_-, \\ C_-' &= i \sin \frac{\alpha}{2} e^{i(\beta-\gamma)/2} C_+ + \cos \frac{\alpha}{2} e^{-i(\beta+\gamma)/2} C_-. \end{aligned} \quad (6.35)$$

Así, partiendo sólo de algunas suposiciones sobre las propiedades del espacio, hemos deducido la transformación de las amplitudes para cualquier rotación. Esto significa que, si conocemos las amplitudes para que cualquier estado de una

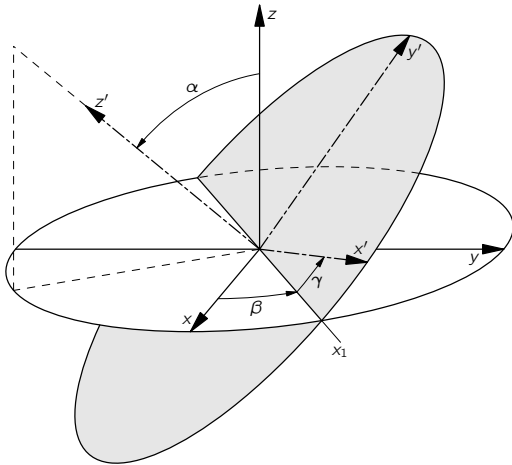


Figura 6-9. La orientación de cualquier sistema de coordenadas x' , y' , z' relativo a otro sistema x , y , z se puede definir en términos de los ángulos de Euler α , β , γ .

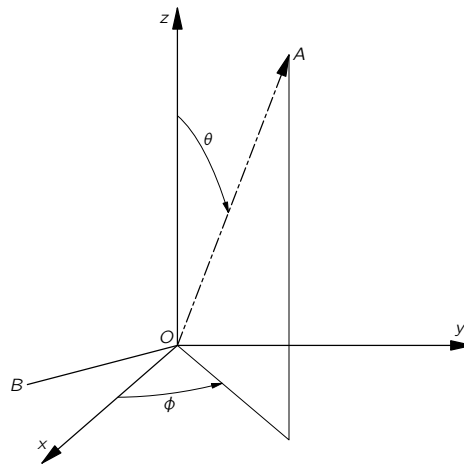


Figura 6-10. Un eje A definido por los ángulos polares θ y ϕ .

[†] Con poco trabajo pueden demostrar que el sistema x , y , z también se puede llevar hacia el sistema x' , y' , z' , por medio de las tres rotaciones siguientes alrededor de los ejes *originales*: (1) roten en un ángulo γ alrededor del eje z original; (2) roten en un ángulo α alrededor del eje x original; (3) roten en un ángulo β alrededor del eje z original.

partícula con espín un medio vaya a los dos haces de un aparato de Stern-Gerlach S , cuyos ejes son x , y y z , podemos calcular qué fracción irá a cada uno de los haces de un aparato T , cuyos ejes son x' , y' y z' . En otras palabras, si tenemos un estado ψ de una partícula con espín un medio, cuyas amplitudes son $C_+ = \langle + | \psi \rangle$ y $C_- = \langle - | \psi \rangle$, para estar “hacia arriba” y “hacia abajo” respecto al eje z de un sistema x, y, z , también conocemos las amplitudes C'_+ y C'_- para estar “hacia arriba” y “hacia abajo” respecto al eje z' de cualquier otro sistema x', y', z' . Los cuatro coeficientes en la ecuación (6.35) son los términos de la “matriz de transformación”, con la cual podemos proyectar las amplitudes de una partícula con espín un medio sobre cualquier otro sistema de coordenadas.

Ahora daremos algunos ejemplos para mostrarles cómo funciona todo esto. Hagamos la siguiente pregunta sencilla. Pongamos un átomo con espín un medio a través de un aparato de Stern-Gerlach que sólo transmite el estado $(+z)$. ¿Cuál es la amplitud para que esté en el estado $(+x)$? El eje $+x$ es lo mismo que el eje $+z'$ de un sistema rotado 90° alrededor del eje y . Para este problema, entonces, lo más simple es utilizar las ecuaciones (6.32) —aunque, por supuesto, pueden utilizar las ecuaciones completas (6.35)—. Como $C_+ = 1$ y $C_- = 0$, obtenemos $C'_+ = 1/\sqrt{2}$. Las probabilidades son el cuadrado absoluto de estas amplitudes; hay un 50% de probabilidad para que la partícula atraviese un aparato que selecciona el estado $(+x)$. Si hubiéramos preguntado sobre el estado $(-x)$, la amplitud habría sido $-1/\sqrt{2}$, que también da una probabilidad $1/2$ —como era de esperar, por la simetría del espacio—. Así, si una partícula está en el estado $(+z)$, es igualmente probable que esté en $(+x)$ o en $(-x)$, pero con fase opuesta.

Tampoco hay problemas con la dirección y . Una partícula en el estado $(+z)$ tiene una posibilidad del 50-50 para estar en $(+y)$ o en $(-y)$. Sin embargo, para éstas (usando la fórmula para una rotación de -90° alrededor de x), las amplitudes son $1/\sqrt{2}$ y $-i/\sqrt{2}$. En este caso, las amplitudes tienen una diferencia de fase de 90° , en vez de la de 180° que tenían para $(+x)$ y $(-x)$. En realidad, es así como se presenta la diferencia entre x y y .

Como ejemplo final, supongan que sabemos que una partícula con espín un medio está en un estado ψ , tal que está polarizada “hacia arriba” respecto a un eje A , definido por los ángulos θ y ϕ en la figura 6-10. Queremos conocer la amplitud C_+ para que la partícula esté “hacia arriba” en la dirección del eje z , y la amplitud C_- para que esté “hacia abajo” en la dirección de z . Podemos encontrar estas amplitudes considerando que A es el eje z de un sistema cuyo eje x está en alguna dirección arbitraria —digamos que en el plano formado por A y z —. Entonces podemos llevar al sistema de A hacia x, y, z mediante tres rotaciones. Primero, hacemos una rotación de $-\pi/2$ alrededor del eje A , la cual lleva al eje x a la línea B en la figura. Luego rotamos θ alrededor de la línea B (nuevo eje x del sistema A) para llevar A hasta el eje z . Finalmente, rotamos en un ángulo $(\pi/2 - \phi)$ alrededor de z . Recordando que sólo tenemos un estado $(+)$ respecto a A , obtenemos

$$C_+ = \cos \frac{\theta}{2} e^{-i\phi/2}, \quad C_- = \sin \frac{\theta}{2} e^{+i\phi/2}. \quad (6.36)$$

Para terminar, nos gustaría resumir los resultados de este capítulo en una forma que sean útiles para nuestro trabajo posterior. En primer lugar, les recordamos que nuestro resultado fundamental en las ecuaciones (6.35) puede escribirse con otra notación. Observen que las ecuaciones (6.35) significan exactamente lo mismo que la ecuación (6.4). Es decir, en las ecuaciones (6.35), los coeficientes de $C_+ = \langle +S | \psi \rangle$ y $C_- = \langle -S | \psi \rangle$ son precisamente las amplitudes $\langle jT | iS \rangle$ de la ecuación (6.4) —las amplitudes para que una partícula que está en el estado i respecto a S también esté en el estado j respecto a T (cuando la orientación de T respecto a S está dada en función de los ángulos α , β y γ)—. También los llamamos R_{ji}^{TS} en la ecuación (6.6). (¡Tenemos una gran cantidad de notaciones!) Por ejemplo, $R_{-+}^{TS} = \langle -T | +S \rangle$ es el coeficiente de C_+ en la fórmula para C'_- , o sea, $i \sin(\alpha/2) e^{i(\beta-\gamma)/2}$. Por consiguiente, podemos resumir nuestros resultados en forma de tabla, como en la tabla 6-1.

De vez en cuando, nos resultará cómodo tener estas amplitudes en forma explícita para algunos casos particulares. Definimos como $R_z(\phi)$ a la rotación en un ángulo ϕ alrededor del eje z . También puede designar a la matriz de rotación correspondiente (omitiendo los subíndices i y j , que están implícitos). Análogamente, $R_x(\phi)$ y $R_y(\phi)$ designarán rotaciones en un ángulo ϕ alrededor del eje x o del eje y . En la tabla 6-2 damos las matrices —las tablas de las amplitudes $\langle jT | iS \rangle$ — que proyectan las amplitudes de un sistema S sobre el sistema T , donde T se obtiene a partir de S por la rotación especificada.

Tabla 6-1

Las amplitudes $\langle jT | iS \rangle$ para una rotación definida por los ángulos de Euler α, β, γ de la figura 6-9

$R_{ji}(\alpha, \beta, \gamma)$		
$\langle jT iS \rangle$	$+S$	$-S$
$+T$	$\cos \frac{\alpha}{2} e^{i(\beta+\gamma)/2}$	$i \operatorname{sen} \frac{\alpha}{2} e^{-i(\beta-\gamma)/2}$
$-T$	$i \operatorname{sen} \frac{\alpha}{2} e^{i(\beta-\gamma)/2}$	$\cos \frac{\alpha}{2} e^{-i(\beta+\gamma)/2}$

Tabla 6-2

Las amplitudes $\langle jT | iS \rangle$ para una rotación $R(\phi)$ en el ángulo ϕ alrededor del eje z , del eje x , o del eje y

$R_z(\phi)$		
$\langle jT iS \rangle$	$+S$	$-S$
$+T$	$e^{i\phi/2}$	0
$-T$	0	$e^{-i\phi/2}$

$R_x(\phi)$		
$\langle jT iS \rangle$	$+S$	$-S$
$+T$	$\cos \phi/2$	$i \operatorname{sen} \phi/2$
$-T$	$i \operatorname{sen} \phi/2$	$\cos \phi/2$

$R_y(\phi)$		
$\langle jT iS \rangle$	$+S$	$-S$
$+T$	$\cos \phi/2$	$\operatorname{sen} \phi/2$
$-T$	$-\operatorname{sen} \phi/2$	$\cos \phi/2$

Dependencia temporal de las amplitudes

7-1 Átomos en reposo, estados estacionarios

Ahora hablaremos un poco acerca del comportamiento de las amplitudes de probabilidad en el tiempo. Decimos un “poco”, porque el comportamiento exacto en el tiempo implica necesariamente el comportamiento en el espacio. Así, entraríamos inmediatamente en la situación más complicada posible, si tuviéramos que hacerlo correctamente y en detalle. Siempre tenemos la dificultad de que podemos tratar algo de un modo lógicamente riguroso, pero bastante abstracto, o en una forma menos rigurosa, pero que nos da cierta idea de una situación real —posponiendo para después el tratamiento más riguroso—. Respecto a la dependencia de la energía, vamos a tomar el segundo camino. Haremos cierto número de suposiciones. No nos esforzaremos en ser rigurosos —sino que solamente les diremos qué se ha encontrado para darles alguna idea acerca del comportamiento de las amplitudes en función del tiempo—. Según vayamos avanzando, aumentará la precisión en la descripción, así que no se pongan nerviosos cuando les parezca que estamos hilando cosas en el aire.

Por supuesto, los hallazgos provienen del aire —del aire en el experimento y en la imaginación de la gente—. Pero nos tomaría mucho tiempo repasar el desarrollo histórico, por lo que tendremos que empezar en alguna parte. Podríamos empezar con lo abstracto y deducir todo —cosa que no entenderían— o podríamos mostrar un gran número de experimentos para justificar cada suposición. Nosotros elegimos un punto intermedio.

Un electrón solo en el espacio vacío puede, bajo ciertas circunstancias, poseer cierta energía definida. Por ejemplo, si está quieto (de modo que no tiene movimiento de traslación, ni momentum, ni energía cinética), tiene su energía de reposo. Un objeto más complicado, como un átomo, también puede tener una energía definida cuando se encuentra quieto, pero también podría ser excitado internamente a otro nivel de energía. (Más tarde describiremos el mecanismo para ello.) Siempre podemos considerar que un átomo en un estado excitado tiene una energía definida, en realidad esto es sólo aproximadamente cierto. Un átomo no permanece excitado para siempre, ya que éste se las arregla para disminuir su energía mediante su interacción con el campo electromagnético. Por lo que hay cierta amplitud para que se genere un estado nuevo —con el átomo en un estado de excitación más bajo, y el campo electromagnético en un estado más alto—. La energía total del sistema es la misma antes y después, pero la energía del *átomo* ha disminuido. Por esto no es preciso decir que un átomo excitado tiene una energía *definida*; pero con frecuencia será conveniente, y no demasiado erróneo, decir sí la tiene.

[Además, ¿por qué las cosas suceden de una manera y no de otra? ¿Por qué un átomo radia luz? La respuesta tiene que ver con la entropía. Cuando la energía está en el campo electromagnético, hay muchas maneras diferentes en las que puede estar —demasiados lugares por donde puede vagar— que si buscamos la condición de equilibrio, encontramos que, en la situación más probable, el campo está excitado con un fotón, y el átomo no está excitado. Se requiere mucho tiempo para que el fotón vuelva y encuentre que puede golpear al átomo, y excitarlo de nuevo. Es muy similar al problema clásico: ¿por qué radia una carga que se acelera? No es porque “quiera” perder energía, ya que, en realidad, cuando ésta radia, la energía del mundo es la misma que antes. La radiación o la absorción van en la dirección en la que crece la *entropía*.]

7-1 Átomos en reposo, estados estacionarios

7-2 Movimiento uniforme

7-3 Energía potencial, conservación de la energía

7-4 Fuerzas: el límite clásico

7-5 La “precesión” de una partícula con espín en un medio

Revisar

Volumen I, capítulo 17,
Espacio-tiempo
Volumen I, capítulo 48,
Pulsaciones

Los núcleos también pueden existir en diferentes niveles de energía, y con una aproximación que desprecie a los efectos electromagnéticos, podemos decir que un núcleo en un estado excitado, permanece en él. Aunque sabemos que no permanece en él para siempre, a veces es útil empezar con una aproximación algo idealizada y más fácil de imaginar. También, muchas veces, es una aproximación adecuada bajo ciertas circunstancias. (Cuando introdujimos por primera vez las leyes clásicas para un cuerpo que cae, no incluimos la fricción, sin embargo, casi nunca encontramos un caso en el que no haya *algo* de fricción.)

Existen las “partículas subnucleares extrañas”, que tienen diversas masas. Pero las más pesadas se desintegran en otras partículas más livianas, por lo que tampoco es correcto decir que tienen una energía exactamente definida. Esto sería cierto sólo si durasen para siempre. Por consiguiente, cuando usamos la aproximación de que tienen una energía definida, estamos olvidando el hecho de que deben desaparecer. Por el momento, desde luego, nos olvidaremos intencionalmente de tales procesos y veremos más adelante cómo tomarlos en cuenta.

Supongamos que tenemos un átomo —o un electrón, o cualquier partícula— que en reposo tiene una energía determinada, E_0 . Por la energía E_0 entendemos la masa de todo el objeto multiplicada por c^2 . Esta masa incluye cualquier energía interna; así que un átomo excitado tiene una masa diferente a la masa del mismo átomo en el estado fundamental. (El *estado fundamental* significa el estado de energía más baja.) Llamaremos E_0 a la “energía en reposo”.

Para un átomo *en reposo*, la *amplitud* cuántica para encontrar a un átomo en un lugar es la *misma para todos los lugares*; ésta *no* depende de la posición. Esto significa, naturalmente, que la *probabilidad* para *encontrar* al átomo en cualquier lugar es la misma. Pero significa todavía más. La *probabilidad* puede ser independiente de la posición, y aun así la *fase* de la *amplitud* puede variar de punto a punto. Pero, para una partícula en reposo, la amplitud completa es igual en todas partes. Sin embargo, sí depende del *tiempo*. Para una partícula en un estado de energía determinada E_0 , la amplitud para encontrar a la partícula en (x, y, z) en el tiempo t es

$$ae^{-i(E_0/\hbar)t}, \quad (7.1)$$

donde a es una constante. La amplitud para estar en cualquier punto en el espacio es la misma para todos los puntos, pero depende del tiempo conforme a (7.1). Simplemente supondremos que esta regla es verdadera.

Naturalmente, podríamos también escribir (7.1) en la forma

$$ae^{-i\omega t}, \quad (7.2)$$

con

$$\hbar\omega = E_0 = Mc^2,$$

donde M es la masa en reposo del estado atómico, o partícula. Hay tres formas diferentes para especificar la energía: por la frecuencia de una amplitud, por la energía en sentido clásico o por la inercia. Todas son equivalentes; sólo son maneras diferentes de decir lo mismo.

Pueden estar pensando que es muy extraño pensar en una “partícula”, la cual tiene amplitudes iguales para estar en todo el espacio. Después de todo, por lo general nos imaginamos a una “partícula” como un objeto pequeño localizado “en un sitio”. Pero no olviden el principio de incertidumbre. Si una partícula tiene una energía determinada, también tiene un momentum determinado. Si la incertidumbre en el momentum es cero, la relación de incertidumbre, $\Delta p \Delta x = \hbar$, nos dice que la incertidumbre en la posición debe ser infinita, y esto es justamente lo que estamos indicando al decir que existe la misma amplitud para encontrar a la partícula en todos los puntos del espacio.

Si las partes internas de un átomo están en diferentes estados, con una energía total diferente, entonces, la variación de la amplitud con el tiempo es diferente. Si no saben en qué estado se encuentra, habrá una cierta amplitud para que esté en un cierto estado, y una cierta amplitud para que esté en otro —y cada una de estas amplitudes tendrá una frecuencia diferente—. Habrá interferencia entre

estas diferentes componentes —como un pulso— la cual puede manifestarse como una probabilidad que varía. Algo estará “ocurriendo” dentro del átomo —aunque se encuentre “en reposo”, en el sentido de que su centro de masa no se mueve—. Sin embargo, si el átomo tiene una energía definida, la amplitud está dada por (7.1), y el cuadrado de la norma de esta amplitud no depende del tiempo. Ven, entonces, que si algo tiene energía definida y hacen alguna pregunta sobre su *probabilidad*, la respuesta es independiente del tiempo. Aunque las *amplitudes* varíen con el tiempo, si la energía está *definida*, variarán como un exponencial imaginario, y la norma no cambiará.

Por esto, decimos con frecuencia que un átomo en un nivel definido de energía está en un *estado estacionario*. Si hacen cualquier medición de las cosas que están dentro, encontrarán que nada (en probabilidad) cambiará con el tiempo. Para que las probabilidades cambien con el tiempo, debemos tener la interferencia de dos amplitudes con frecuencias diferentes, y esto significa que no podemos saber cuál es la energía. El objeto tendrá una amplitud para estar en un estado de una energía, y otra amplitud para estar en un estado de otra energía. Ésta es la descripción cuántica de algo cuando su *comportamiento* depende del tiempo.

Si tenemos una “condición” que es una mezcla de dos estados distintos con energías diferentes, entonces la amplitud para cada uno de los dos estados varía con el tiempo conforme a la ecuación (7.2), por ejemplo, como

$$e^{-i(E_1/\hbar)t} \quad \text{y} \quad e^{-i(E_2/\hbar)t}. \quad (7.3)$$

Y si tenemos alguna combinación de las dos, tendremos una interferencia. Pero noten que si añadimos una constante a ambas energías, no cambia nada. Si alguien más usa una escala diferente de energía, en la cual todas las energías fueran aumentadas (o disminuidas) en una cantidad constante —digamos, en la cantidad A —, entonces las amplitudes en los dos estados, desde este punto de vista, serían

$$e^{-i(E_1+A)t/\hbar} \quad \text{y} \quad e^{-i(E_2+A)t/\hbar}. \quad (7.4)$$

Todas sus amplitudes estarían multiplicadas por el mismo factor $e^{-i(A/\hbar)t}$, y todas las combinaciones lineales, o interferencias, tendrían el mismo factor. Al calcular el cuadrado de las normas para encontrar las probabilidades, todas las respuestas serían iguales. La elección de un origen para nuestra escala de energía no importa; podemos medir energías a partir del origen que queramos. Para cuestiones relativistas, es cómodo medir la energía de manera que se incluya la masa en reposo, pero para otros fines que no son relativistas, también es cómodo restar alguna cantidad estándar de todas las energías que aparezcan. Por ejemplo, en el caso de un átomo, generalmente conviene restar la energía $M_s c^2$, donde M_s es la masa de todas las piezas por *separado* —el núcleo y los electrones—, que es, naturalmente, diferente de la masa del átomo. En otros problemas, puede ser útil restar a todas las energías la cantidad $M_g c^2$, donde M_g es la masa de todo el átomo en el *estado fundamental*; así, la energía que aparece es justamente la energía de excitación del átomo. De este modo, algunas veces podremos mover nuestro cero de energía por una constante muy grande, pero no importa, siempre que movamos todas las energías en una situación determinada por la misma constante. Hasta aquí con una partícula que está quieta.

7-2 Movimiento uniforme

Si suponemos que la teoría de la relatividad es correcta, una partícula en reposo en un sistema inercial puede estar en movimiento uniforme en otro sistema inercial. En el sistema en reposo de la partícula, la amplitud de probabilidad es la misma para todas x , y y z , pero varía con t . La *magnitud* de la amplitud es la misma para toda t , pero la *fase* depende de t . Podemos obtener una imagen del comportamiento de la amplitud si graficamos líneas de igual fase —por ejemplo, líneas de fase cero— como función de x y de t . Para una partícula en reposo, estas líneas de fase igual son paralelas al eje x , y están igualmente espaciadas en la coordenada t , como se muestra con las líneas discontinuas en la figura 7-1.

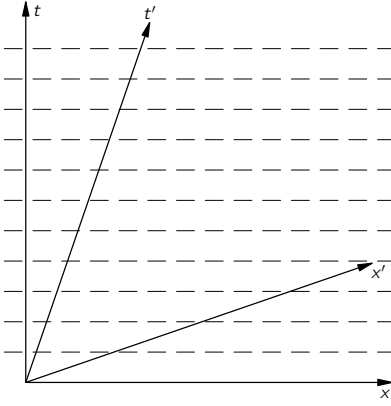


Figura 7-1. Transformación relativista de la amplitud para una partícula en reposo en los sistemas x - t .

En un sistema diferente — x' , y' , z' y t' — que se mueve respecto a la partícula en la dirección x , por ejemplo, las coordenadas x' y t' , de cualquier punto determinado en el espacio, están relacionadas con x y t por la transformación de Lorentz. Esta transformación se puede representar gráficamente dibujando los ejes x' y t' , como se ha hecho en la figura 7-1. (Véase capítulo 17 del volumen I, figura 17-2.) Pueden observar que, en el sistema x' - t' , los puntos de fase igual† tienen un espaciamiento diferente a lo largo del eje t' , por lo que la frecuencia de la variación temporal es diferente. También existe una variación de la fase con x' , por lo que la amplitud de probabilidad debe ser una función de x' .

Usando la transformación de Lorentz para la velocidad v , por ejemplo, a lo largo de la dirección negativa del eje x , el tiempo t está relacionado con el tiempo t' mediante:

$$t = \frac{t' - x'v/c^2}{\sqrt{1 - v^2/c^2}},$$

por lo que nuestra amplitud ahora varía como

$$e^{-(i/\hbar)E_0 t} = e^{-(i/\hbar)(E_0 t' / \sqrt{1 - v^2/c^2} - E_0 v x' / c^2 \sqrt{1 - v^2/c^2})}.$$

En el sistema con primas, la amplitud varía tanto en el espacio como en el tiempo. Si escribimos la amplitud en la forma

$$e^{-(i/\hbar)(E'_p t' - p' x')},$$

vemos que $E'_p = E_0 / \sqrt{1 - v^2/c^2}$ es la energía, que se calcula en la manera clásica, para una partícula con una energía en reposo E_0 que viaja a la velocidad v , y que $p' = E'_p v / c^2$ es el correspondiente momentum de la partícula.

Ustedes saben que $x_\mu = (t, x, y, z)$ y $p_\mu = (E, p_x, p_y, p_z)$ son vectores con cuatro componentes, y que $p_\mu x_\mu = Et - \mathbf{p} \cdot \mathbf{x}$ es un invariante escalar. En el sistema de referencia en reposo de la partícula, $p_\mu x_\mu$ es igual a Et ; así que si transformamos a otro sistema de referencia, Et es remplazado por

$$E' t' - \mathbf{p}' \cdot \mathbf{x}'.$$

De este modo, la amplitud de probabilidad para una partícula, que tiene el momentum $\mathbf{p}$, será proporcional a

$$e^{-(i/\hbar)(E_p t - \mathbf{p} \cdot \mathbf{x})}, \quad (7.5)$$

donde E_p es la energía de la partícula cuyo momentum es p , esto es,

$$E_p = \sqrt{(pc)^2 + E_0^2}, \quad (7.6)$$

donde E_0 es, como antes, la energía en reposo. Para problemas no relativistas, podemos escribir

$$E_p = M_s c^2 + W_p, \quad (7.7)$$

donde W_p es la energía por encima de la energía en reposo $M_s c^2$ de las partes del átomo. En general, W_p incluye tanto a la energía cinética del átomo, así como su energía de enlace o de excitación, a la que llamaremos energía “interna”. Escribimos:

$$W_p = W_{\text{int}} + \frac{p^2}{2M}, \quad (7.8)$$

y las amplitudes serían

$$e^{-(i/\hbar)(W_p t - \mathbf{p} \cdot \mathbf{x})}. \quad (7.9)$$

Ya que trataremos, por lo general, con cálculos no relativistas, usaremos esta forma para las amplitudes de probabilidad.

† Estamos suponiendo que la fase debe tener el mismo valor para puntos correspondientes en los dos sistemas. Este punto es muy delicado, ya que la fase de la amplitud cuántica es, en gran medida, arbitraria. Una justificación completa de esta suposición requiere una discusión más detallada, que involucra interferencias de dos o más amplitudes.

Noten que nuestra transformación relativista nos ha dado la variación de la amplitud para un átomo que se mueve en el espacio sin suposiciones adicionales. El número de onda para las variaciones en el espacio es, de acuerdo con (7.9),

$$k = \frac{p}{\hbar}; \quad (7.10)$$

por lo que la longitud de onda es

$$\lambda = \frac{2\pi}{k} = \frac{h}{p}. \quad (7.11)$$

Ésta es la misma longitud de onda que usamos antes para partículas con momentum p . El primero que obtuvo esta fórmula fue De Broglie, exactamente de la misma manera. Para una partícula en movimiento, la *frecuencia* de las variaciones en la amplitud está dada por

$$\hbar\omega = W_p. \quad (7.12)$$

El cuadrado de la norma de (7.9) es exactamente 1, de ahí que, para una partícula en movimiento con una *energía definida*, la probabilidad de encontrarla es la misma para todos los puntos y no cambia con el tiempo. (Es importante notar que la amplitud es una onda *compleja*. Si usáramos una onda sinusoidal real, la norma variaría de punto a punto, lo que sería incorrecto.)

Sabemos, naturalmente, que hay situaciones en las que las partículas se mueven de un lugar a otro, por lo que la probabilidad depende de la posición y cambia con el tiempo. ¿Cómo describimos tales situaciones? Podemos hacerlo considerando amplitudes que sean una superposición de dos o más amplitudes correspondientes a estados de energía definida. Discutimos esta situación en el capítulo 48 del volumen I —¡hasta para amplitudes de probabilidad!—. Encontramos que la suma de dos amplitudes con diferentes números de onda k (es decir, momentum) y frecuencias ω (es decir, energías) nos daba curvas de interferencia, o pulsaciones, de tal modo que el cuadrado de la amplitud variaba en el espacio y en el tiempo. También encontramos que estas pulsaciones se mueven con la llamada “velocidad de grupo”, dada por:

$$v_g = \frac{\Delta\omega}{\Delta k},$$

donde Δk y $\Delta\omega$ son las diferencias entre los números de onda y las frecuencias de las dos ondas. Para ondas más complicadas —formadas por la suma de muchas amplitudes de frecuencias muy próximas— la velocidad de grupo es

$$v_g = \frac{d\omega}{dk}. \quad (7.13)$$

Usando $\omega = E_p/\hbar$ y $k = p/\hbar$, vemos que

$$v_g = \frac{dE_p}{dp}. \quad (7.14)$$

Usando la ecuación (7.6), tenemos que

$$\frac{dE_p}{dp} = c^2 \frac{p}{E_p}. \quad (7.15)$$

Pero $E_p = Mc^2$, de donde

$$\frac{dE_p}{dp} = \frac{p}{M}, \quad (7.16)$$

que es justamente la velocidad clásica de la partícula. De manera alternativa, si usamos las expresiones no relativistas, tenemos

$$\omega = \frac{W_p}{\hbar} \quad \text{y} \quad k = \frac{p}{\hbar},$$

y

$$\frac{d\omega}{dk} = \frac{dW_p}{dp} = \frac{d}{dp} \left(\frac{p^2}{2M} \right) = \frac{p}{M}, \quad (7.17)$$

que es, de nuevo, la velocidad clásica.

Nuestro resultado es que, si tenemos varias amplitudes para estados puros de energía de casi la misma energía, su interferencia nos da “paquetes” en la probabilidad, que se mueven en el espacio con una velocidad igual a la velocidad de una partícula clásica de la misma energía. Deberíamos señalar que, sin embargo, cuando decimos que podemos sumar dos amplitudes de diferentes números de onda, para obtener una pulsación que correspondería a una partícula en movimiento, hemos introducido algo nuevo —algo que no podemos deducir a partir de la teoría de la relatividad—. Dijimos cómo se comportaba la amplitud para una partícula que está quieta, y luego dedujimos lo que haría si la partícula estuviese en movimiento. Pero *no podemos* deducir a partir de estos argumentos lo que sucedería cuando hay *dos* ondas que se mueven con diferentes velocidades. Si detenemos a una, no podemos detener a la otra. Así que, tácitamente, hemos agregado la hipótesis *adicional* de que no solamente (7.9) es una solución *posible*, sino de que también puede haber soluciones con toda clase de p para el mismo sistema, y que los términos diferentes interferirán.

7-3 Energía potencial, conservación de la energía

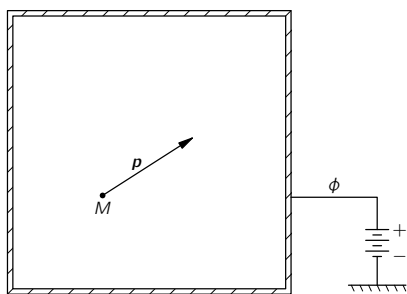


Figura 7-2. Una partícula de masa M y momentum $\mathbf{p}$ en una región de potencial constante.

Ahora nos gustaría discutir lo que sucede cuando la energía de una partícula puede cambiar. Comencemos considerando una partícula que se mueve en un campo de fuerzas descrito por un potencial. Discutiremos primero el efecto de un potencial constante. Supongamos que tenemos una caja grande de metal, a la que le hemos dado cierto potencial electrostático ϕ , como se muestra en la figura 7-2. Si hay objetos cargados dentro de la caja, su energía potencial será $q\phi$, a la que llamaremos V , y será por completo independiente de la posición. Entonces no puede haber ningún cambio físico adentro, porque el potencial constante no hace ninguna diferencia en lo que ocurre dentro de la caja. Ahora, no hay ningún modo de deducir cuál sería la respuesta, por lo que debemos hacer una conjetura. La conjetura apropiada es, más o menos, la que están esperando: para la energía, debemos usar la suma de la energía potencial V y la energía E_p —que es en sí la suma de las energías interna y cinética—. La amplitud es proporcional a

$$e^{-(i/\hbar)[(E_p+V)t-\mathbf{p}\cdot\mathbf{x}]} \quad (7.18)$$

El *principio general* es que el coeficiente de t , que podemos llamar ω , siempre está dado por la *energía total* del sistema: energía interna (o “música”), más energía cinética, más energía potencial:

$$\hbar\omega = E_p + V. \quad (7.19)$$

O, para situaciones no relativistas,

$$\hbar\omega = W_{\text{int}} + \frac{p^2}{2M} + V. \quad (7.20)$$

Ahora bien, ¿qué pasa con los fenómenos físicos dentro de la caja? Si hay varios estados de energía diferentes, ¿qué obtendremos? La amplitud para cada estado tiene el mismo factor adicional

$$e^{-(i/\hbar)Vt}$$

además del que tuviera con $V = 0$. Esto es exactamente como si hubiésemos cambiado el cero en nuestra escala de energía. Produce el mismo cambio de fase en todas las amplitudes, pero como hemos visto antes, esto no cambia ninguna de las probabilidades. Todos los fenómenos físicos siguen iguales. (Hemos supuesto que estamos hablando de estados diferentes del mismo objeto cargado, así que $q\phi$ es igual para todos. Si un objeto pudiese cambiar su carga al ir de un estado a otro, tendríamos un resultado diferente, pero la conservación de la carga impide que esto ocurra.)

Hasta aquí, nuestra suposición concuerda con lo que esperaríamos para un cambio en el nivel de energía de referencia. Pero si es realmente correcto, debe

servir para una energía potencial que no es simplemente una constante. En general, V puede variar de un modo arbitrario con el tiempo y el espacio, y el resultado completo para la amplitud se debe dar en términos de una ecuación diferencial. No queremos preocuparnos con el caso general por ahora, sino que sólo queremos tener una idea de cómo suceden las cosas, por eso pensaremos únicamente en un potencial que es constante en el tiempo y que varía muy lentamente en el espacio. Así, podemos hacer una comparación entre las ideas clásicas y cuánticas.

Supongan que consideramos la situación de la figura 7-3, en la que tenemos dos cajas con potenciales constantes ϕ_1 y ϕ_2 , y una región entre ellas, donde supondremos que el potencial varía suavemente de una a otra. Imaginemos que una partícula tiene cierta amplitud para encontrarse en cualquiera de estas regiones. También suponemos que el momentum es lo bastante grande como para que, en cualquier región pequeña, donde haya muchas longitudes de onda, el potencial sea casi constante. En consecuencia, pensaríamos que, en cualquier parte del espacio, la amplitud se debería parecer a (7.18) con la V apropiada para esa región del espacio.

Pensemos en el caso especial en que $\phi_1 = 0$, por lo que la energía potencial ahí es cero, pero que $q\phi_2$ es negativo, de manera que, clásicamente, la partícula tendría más energía en la segunda caja. Clásicamente, se movería más rápido en la segunda caja —tendría más energía y, por lo tanto, más momentum—. Veamos cómo ocurriría esto de la mecánica cuántica.

Según nuestra suposición, la amplitud en la primera caja sería proporcional a

$$e^{-(i/\hbar)[(W_{\text{int}} + p_1^2/2M + V_1)t - \mathbf{p}_1 \cdot \mathbf{x}]}, \quad (7.21)$$

y la amplitud en la segunda caja sería proporcional a

$$e^{-(i/\hbar)[(W_{\text{int}} + p_2^2/2M + V_2)t - \mathbf{p}_2 \cdot \mathbf{x}]}. \quad (7.22)$$

(Digamos que la energía interna no está variando, sino que permanece igual en ambas regiones.) La pregunta es: ¿cómo estas dos amplitudes se unen a través de la región entre las dos cajas?

Vamos a suponer que todos los potenciales son constantes en el tiempo —de modo que nada varíe en las condiciones—. También supondremos que las variaciones en la amplitud (es decir, su fase) tienen la misma *frecuencia* en todas partes —porque, por así decirlo, no hay nada en el “medio” que dependa del tiempo—. Si nada cambia en el espacio, podemos considerar que la onda en una región “genera” ondas subsidiarias en todo el espacio, que oscilarán con la misma frecuencia —igual que las ondas de luz, que atraviesan materiales en reposo, no cambian su frecuencia—. Si las frecuencias en (7.21) y (7.22) son las mismas, debemos tener

$$W_{\text{int}} + \frac{p_1^2}{2M} + V_1 = W_{\text{int}} + \frac{p_2^2}{2M} + V_2. \quad (7.23)$$

Ambos lados son justo las energías totales clásicas, por eso la ecuación (7.23) es un enunciado de la conservación de la energía. En otras palabras, el enunciado clásico de la conservación de la energía es equivalente al enunciado cuántico de que las frecuencias para una partícula son las mismas en todas partes, si las condiciones no cambian con el tiempo. Todo concuerda con la idea de que $\hbar\omega = E$.

En el ejemplo especial en que $V_1 = 0$ y V_2 es negativo, la ecuación (7.23) nos da que p_2 es mayor que p_1 , por lo que la longitud de las ondas es más corta en la región 2. Las superficies con la misma fase se muestran con líneas discontinuas en la figura 7-3. También hemos dibujado una gráfica de la parte real de la amplitud, que nos muestra de nuevo cómo decrece la longitud de onda al pasar de la región 1 a la 2. La velocidad de grupo de las ondas, que es p/M , también aumenta en la forma que esperaríamos a partir de la conservación de la energía clásica, ya que es exactamente igual a la ecuación (7.23).

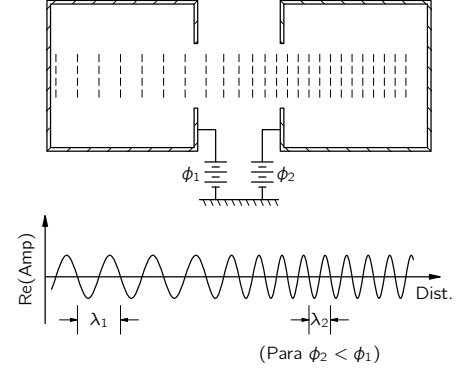


Figura 7-3. La amplitud para una partícula en movimiento de un potencial a otro.

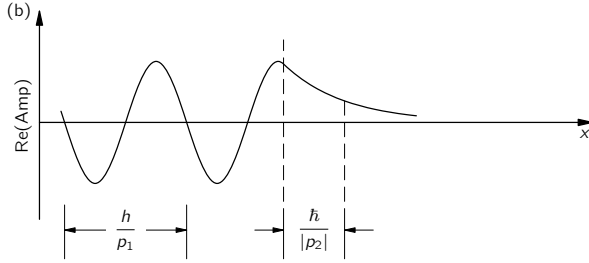
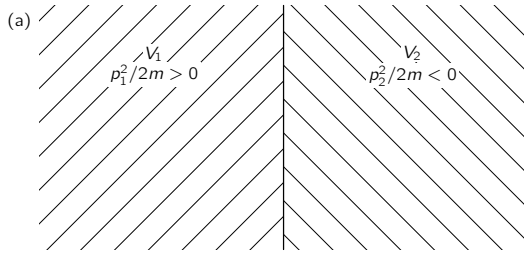


Figura 7-4. Amplitud para una partícula que se acerca a un potencial altamente repulsivo.

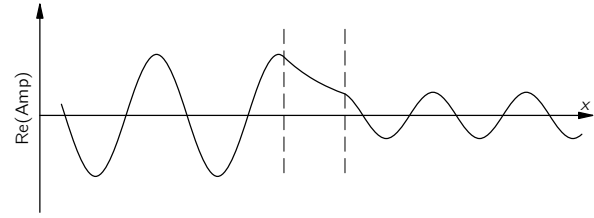
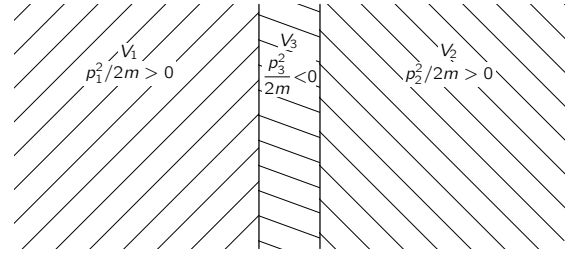


Figura 7-5. Penetración de la amplitud a través de una barrera de potencial.

Hay un caso especial interesante cuando V_2 se hace tan grande que $V_2 - V_1$ es más grande que $p_1^2/2M$. Entonces p_2^2 , que está dado por

$$p_2^2 = 2M \left[\frac{p_1^2}{2M} - V_2 + V_1 \right], \quad (7.24)$$

es *negativo*. Esto significa que p_2 es un número imaginario, digamos ip' . De manera clásica, diríamos que la partícula nunca entra en la región 2 —no tiene energía suficiente para subir el escalón de potencial—. De manera cuántica, sin embargo, la amplitud sigue estando dada por la ecuación (7.22); su variación en el espacio sigue siendo

$$e^{(i/\hbar)p_2 \cdot x}.$$

Pero si p_2 es imaginario, la dependencia espacial se transforma en una exponencial real. Digamos que la partícula se movía inicialmente en la dirección $+x$, entonces la amplitud varía como

$$e^{-p'x/\hbar}. \quad (7.25)$$

La amplitud decrece rápidamente al aumentar x .

Imaginen que las dos regiones con potenciales diferentes se encontrasen muy próximas, de modo que la energía potencial cambiase de manera repentina de V_1 a V_2 , como lo muestra en la figura 7-4(a). Si graficamos la parte real de la amplitud de probabilidad, obtenemos la dependencia mostrada en la parte (b) de la misma figura. La onda en la primera región corresponde a una partícula tratando de entrar en la segunda región, pero la amplitud cae ahí rápidamente. Hay cierta probabilidad para que sea observada en la segunda región —donde *nunca* podría llegar clásicamente— pero la amplitud es muy pequeña, excepto muy cerca del borde. La situación es bastante parecida a la que encontramos para la reflexión interna total de la luz. Normalmente la luz no sale, pero podemos observarla si colocamos algo a una distancia de una o dos longitudes de onda de la superficie.

Recordarán que, si colocamos una segunda superficie cerca del borde donde la luz era reflejada totalmente, podíamos obtener un poco de luz transmitida hacia el segundo pedazo de material. Algo semejante sucede con las partículas en la mecánica cuántica. Si hay una región estrecha con un potencial V , tan grande que la energía cinética clásica fuese negativa, la partícula no pasaría nunca clásicamente. Pero cuánticamente, la amplitud, que decae exponencialmente,

puede llegar a la región y dar una pequeña probabilidad para que la partícula se encuentre en el otro lado, donde la energía cinética es de nuevo positiva. La situación se ilustra en la figura 7-5. Este efecto se llama “penetración cuántica de una barrera”.

La penetración de una barrera por una amplitud cuántica da la explicación —o descripción— del decaimiento α del núcleo de uranio. La energía potencial de una partícula α , en función de su distancia al centro, se muestra en la figura 7-6(a). Si tratamos de disparar una partícula α con una energía E hacia el núcleo, sentiría una repulsión electrostática debida a la carga nuclear z y, clásicamente, no llegaría más cerca que la distancia r_1 , donde su energía total es igual a la energía potencial V . Sin embargo, más cerca la energía potencial es mucho menor, a causa de la fuerte atracción de las fuerzas nucleares de corto alcance. ¿Cómo es, entonces, que en la desintegración radiactiva encontramos partículas α , que proceden del interior de un núcleo y salen con la energía E ? Porque comienzan dentro del núcleo con la energía E , y “se filtran” a través de la barrera de potencial. La amplitud de probabilidad es más o menos como se muestra en la parte (b) de la figura 7-6, aunque, en realidad, la caída exponencial es más grande de lo que se muestra. Es realmente notable que la vida media de una partícula α en el núcleo de uranio sea tan grande como $4\frac{1}{2}$ miles de millones de años, cuando las oscilaciones naturales dentro del núcleo son extremadamente rápidas —¡alrededor de 10^{22} por segundo!—. ¿Cómo podemos obtener un número como 10^9 años a partir de 10^{-22} segundos? La respuesta es que el exponencial provoca el factor tremendamente pequeño de e^{-45} —que da la tan pequeña, aunque definida, probabilidad para filtrarse—. Una vez que la partícula α está en el núcleo, no hay casi ninguna probabilidad para encontrarla afuera; sin embargo, si toman muchos núcleos y esperan lo suficiente, pueden tener suerte y encontrar una que haya salido.

7-4 Fuerzas: el límite clásico

Supongan que tenemos una partícula que se mueve y atraviesa una región donde hay un potencial, el cual varía perpendicularmente a la dirección del movimiento. Clásicamente, describiríamos la situación como se muestra en la figura 7-7. Si la partícula se mueve en la dirección de x , y entra en una región donde hay un potencial que varía con y , la partícula sufrirá una aceleración transversal debido a la fuerza $F = -\partial V/\partial y$. Si la fuerza está presente solamente en una región limitada de ancho w , la fuerza actuará sólo durante el tiempo w/v . A la partícula se le dará el momentum transversal

$$p_y = F \frac{w}{v}.$$

El ángulo de desviación $\delta\theta$ es entonces

$$\delta\theta = \frac{p_y}{p} = \frac{Fw}{pv},$$

donde p es el momentum inicial. Usando $-\partial V/\partial y$ en lugar de F , obtenemos

$$\delta\theta = -\frac{w}{pv} \frac{\partial V}{\partial y}. \quad (7.26)$$

Ahora nos corresponde verificar si nuestra idea de que las ondas cumplen con (7.20) puede explicar el mismo resultado. Examinemos la misma situación cuánticamente, suponiendo que todo está en una escala muy grande comparada con la longitud de onda de nuestras amplitudes de probabilidad. En cualquier región pequeña podemos decir que la amplitud varía así

$$e^{-(i/\hbar)[(W+p^2/2M+V)t-\mathbf{p}\cdot\mathbf{x}]} \quad (7.27)$$

¿Podremos decir que esto también originará una desviación de la partícula cuando V tenga un gradiente transversal? Hemos bosquejado en la figura 7-8 cómo se

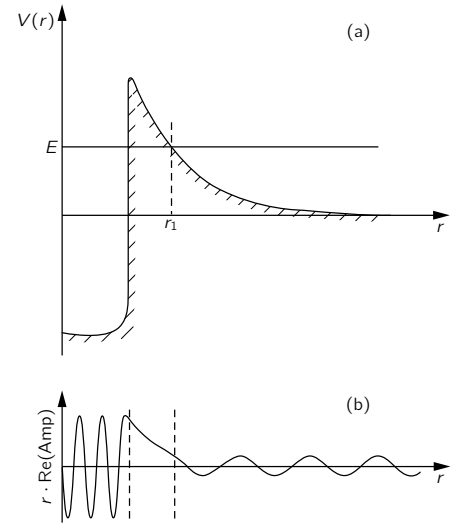


Figura 7-6. (a) La función potencial para una partícula α en un núcleo de uranio. (b) Forma cualitativa de la amplitud de probabilidad.

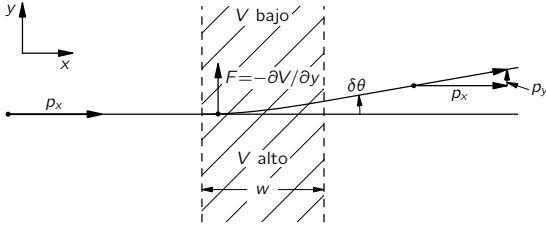


Figura 7-7. Desviación de una partícula mediante un gradiente de potencial transversal.

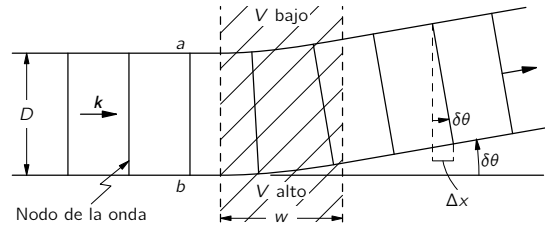


Figura 7-8. La amplitud de probabilidad en una región donde hay un gradiente de potencial transversal.

verán las ondas de amplitud de probabilidad. Hemos trazado un conjunto de “nodos de onda”, que consideramos como superficies donde la fase de la amplitud es cero. En cada región muy pequeña, la longitud de onda —la distancia entre nodos sucesivos— es

$$\lambda = \frac{h}{p},$$

donde p se relaciona con V mediante

$$W + \frac{p^2}{2M} + V = \text{Constante.} \quad (7.28)$$

En la región donde V es más grande, p es más pequeño y la longitud de onda es mayor. Por lo que el ángulo de los nodos de las ondas cambia como se muestra en la figura.

Para encontrar la variación en el ángulo de los nodos de la onda notamos que, para los dos caminos a y b de la figura 7-8, hay una diferencia de potencial $\Delta V = (\partial V / \partial y) D$, de modo que existe una diferencia Δp en el momentum a lo largo de los dos caminos, la cual se puede obtener de (7.28):

$$\Delta \left(\frac{p^2}{2M} \right) = \frac{p}{M} \Delta p = -\Delta V. \quad (7.29)$$

El número de onda $p/\hbar$ es, por lo tanto, diferente a lo largo de los dos caminos, lo que significa que la fase está avanzando a una rapidez diferente. La diferencia en la rapidez de aumento de fase es $\Delta k = \Delta p / \hbar$, de modo que la diferencia de fase acumulada en la distancia total w es

$$\Delta(\text{fase}) = \Delta k \cdot w = \frac{\Delta p}{\hbar} \cdot w = -\frac{M}{p\hbar} \Delta V \cdot w. \quad (7.30)$$

Ésta es la cantidad por la cual la fase en el camino b se “adelanta” de la fase en el camino a cuando la onda abandona la zona. Pero fuera de la zona, una fase que se adelanta en esta cantidad corresponde a un nodo de la onda que se encuentra adelantado en la cantidad

$$\Delta x = \frac{\lambda}{2\pi} \Delta(\text{fase}) = \frac{\hbar}{p} \Delta(\text{fase})$$

o

$$\Delta x = -\frac{M}{p^2} \Delta V \cdot w. \quad (7.31)$$

Refiriéndonos a la figura 7-8, vemos que los nuevos frentes de onda se encontrarán en un ángulo $\delta\theta$ dado por

$$\Delta x = D \delta\theta; \quad (7.32)$$

por lo que tenemos

$$D \delta\theta = -\frac{M}{p^2} \Delta V \cdot w. \quad (7.33)$$

Que es igual a la ecuación (7.26) si sustituimos p/M por v y $\Delta V/D$ por $\partial V / \partial y$.

El resultado que hemos obtenido es correcto sólo si las variaciones en el potencial son lentas y suaves —lo que llamamos el *límite clásico*—. Hemos mostrado que bajo estas condiciones, obtenemos los mismos movimientos para una partícula que los que obtenemos de $F = ma$, siempre que supongamos que un potencial contribuye a la amplitud de probabilidad con una fase igual a $Vt/\hbar$. *En el límite clásico, la mecánica cuántica está de acuerdo con la mecánica newtoniana.*

7-5 La “precesión” de una partícula con espín un medio

Noten que no hemos supuesto nada especial sobre la energía potencial —sólo que es aquella energía cuya derivada nos da una fuerza—. Por ejemplo, en el experimento de Stern-Gerlach teníamos la energía $U = -\boldsymbol{\mu} \cdot \mathbf{B}$, que nos da una fuerza si $\mathbf{B}$ tiene una variación espacial. Si quisiéramos dar una descripción cuántica, hubiéramos dicho que las partículas en un haz tenían una energía que variaba de una manera, mientras que las de otro haz tenían una variación de energía opuesta. (Podríamos incluir a la energía magnética U dentro de la energía potencial V , o dentro de la energía “interna” W ; eso no importa.) A causa de la variación de energía, las ondas se refractan, y los haces se desvían hacia arriba o hacia abajo. (Ahora sabemos que la mecánica cuántica nos dará la misma desviación que calculamos con la mecánica clásica.)

A partir de la dependencia de la amplitud con la energía potencial, también podríamos esperar que si una partícula está en un campo magnético uniforme en la dirección z , su amplitud de probabilidad debe estar cambiando con el tiempo de acuerdo a

$$e^{-(i/\hbar)(-\mu_z B)t}.$$

(Podemos considerar que esto es, en efecto, una definición de μ_z .) En otras palabras, si colocamos una partícula en un campo uniforme B durante un tiempo τ , su amplitud de probabilidad será multiplicada por

$$e^{-(i/\hbar)(-\mu_z B)\tau}$$

respecto a lo que sería en ausencia de campo. Ya que para una partícula con espín un medio, μ_z puede ser más o menos algún número, digamos μ , los dos estados posibles en un campo uniforme tendrían sus fases cambiando a la misma rapidez, pero en direcciones opuestas. Las dos amplitudes están multiplicadas por

$$e^{\pm(i/\hbar)\mu B\tau}. \quad (7.34)$$

Este resultado tiene algunas consecuencias interesantes. Supongan que tenemos una partícula con espín un medio en algún estado, el cual no es de espín hacia arriba puro o de espín hacia abajo puro. Podemos describir su estado en términos de las amplitudes para estar en un estado hacia arriba puro o hacia abajo puro. Pero en un campo magnético, estos dos estados tendrán fases que cambian a diferentes rapidezces. Así que si preguntamos algo acerca de las amplitudes, la respuesta dependerá del tiempo que hayan estado en el campo.

Como ejemplo, consideramos la desintegración del muon en un campo magnético. Cuando se producen muones como productos de desintegración de mesones π , éstos están polarizados (en otras palabras, tienen una dirección preferida para el espín). Los muones, a su vez, se desintegran —después de unos 2.2 microsegundos en promedio— emitiendo un electrón y dos neutrinos:

$$\mu \rightarrow e + \nu + \bar{\nu}.$$

En esta desintegración ocurre que (al menos para las energías más altas) los electrones son emitidos preferentemente en la dirección opuesta a la dirección del espín del muon.

Supongan entonces que consideramos el arreglo experimental mostrado en la figura 7-9. Si los muones polarizados entran desde la izquierda y se los lleva al reposo en un bloque de material en A , un poco más tarde, se desintegrarán.

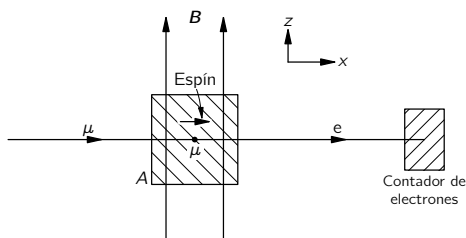


Fig. 7-9. Un experimento de desintegración de un muón.

Los electrones emitidos saldrán, en general, en todas las direcciones posibles. Supongan, sin embargo, que todos los muones entran en el bloque de detención en A con sus espines en la dirección x . Sin un campo magnético, habrá cierta distribución angular de direcciones de desintegración; nos gustaría saber cómo cambia esta distribución por la presencia de un campo magnético. Esperamos que varíe de alguna manera con el tiempo. Podemos descubrir lo que sucede preguntando, para cualquier momento, cuál es la amplitud para que el muon se encuentre en el estado $(+x)$.

Podemos presentar el problema en la siguiente manera. Se sabe que un muon tiene su espín en la dirección $+x$ en el tiempo $t = 0$, ¿cuál es la amplitud para que se encuentre en el mismo estado en el tiempo τ ? Ahora no tenemos ninguna regla para el comportamiento de una partícula con espín un medio en un campo magnético perpendicular al espín, pero sí sabemos lo que le pasa a los estados de espín hacia arriba y de espín hacia abajo respecto al campo —sus amplitudes se multiplican por el factor (7.34)—. Nuestro procedimiento consiste, entonces, en escoger la representación en la que los estados base son de espín hacia arriba y hacia abajo respecto a la dirección z (la dirección del campo). Cualquier pregunta se puede expresar, entonces, haciendo referencia a las amplitudes para estos estados.

Digamos que $\psi(t)$ representa al estado del muon. Cuando entra en el bloque A , su estado es $\psi(0)$, y queremos conocer $\psi(\tau)$ en un tiempo posterior τ . Si representamos a los dos estados base con $(+z)$ y $(-z)$, conocemos las dos amplitudes $\langle +z | \psi(0) \rangle$ y $\langle -z | \psi(0) \rangle$ —conocemos estas amplitudes porque sabemos que $\psi(0)$ representa un estado con el espín en el estado $(+x)$ —. De acuerdo con los resultados del capítulo anterior, estas amplitudes son†

$$\langle +z | +x \rangle = C_+ = \frac{1}{\sqrt{2}}$$

y

$$\langle -z | +x \rangle = C_- = \frac{1}{\sqrt{2}}.$$

Sucede que son iguales. Puesto que estas amplitudes se refieren a la condición en $t = 0$, llamémoslas $C_+(0)$ y $C_-(0)$.

Ahora sabemos lo que sucede a estas dos amplitudes con el tiempo. Usando (7.34), tenemos

$$C_+(t) = C_+(0)e^{-(i/\hbar)\mu Bt}$$

y

$$C_-(t) = C_-(0)e^{+(i/\hbar)\mu Bt}.$$

Pero si conocemos $C_+(t)$ y $C_-(t)$, tenemos todo lo que hay que saber acerca de la condición para t . La única dificultad es que *queremos* saber cuál es la probabilidad para que en t el espín esté en la dirección $+x$. Nuestras reglas generales pueden, sin embargo, resolver este problema. Decimos que la amplitud para estar en el estado $(+x)$ en el tiempo t , que llamaremos $A_+(t)$, es

$$A_+(t) = \langle +x | \psi(t) \rangle = \langle +x | +z \rangle \langle +z | \psi(t) \rangle + \langle +x | -z \rangle \langle -z | \psi(t) \rangle$$

o

$$A_+(t) = \langle +x | +z \rangle C_+(t) + \langle +x | -z \rangle C_-(t).$$

Usando de nuevo los resultados del capítulo anterior —o mejor la igualdad $\langle \phi | \chi \rangle = \langle \chi | \phi \rangle^*$ del capítulo 5— sabemos que

$$\langle +x | +z \rangle = \frac{1}{\sqrt{2}}, \quad \langle +x | -z \rangle = \frac{1}{\sqrt{2}}.$$

† Si omitieron el capítulo 6, pueden tomar a (7.35) como una regla que no hemos deducido por ahora. Más adelante (en el capítulo 10) discutiremos en detalle la precesión del espín, incluyendo una derivación de estas amplitudes.

Por lo que conocemos todas las cantidades de la ecuación (7.37). Obtenemos

$$A_+(t) = \frac{1}{2}e^{(i/\hbar)\mu Bt} + \frac{1}{2}e^{-(i/\hbar)\mu Bt},$$

o

$$A_+(t) = \cos \frac{\mu B}{\hbar} t.$$

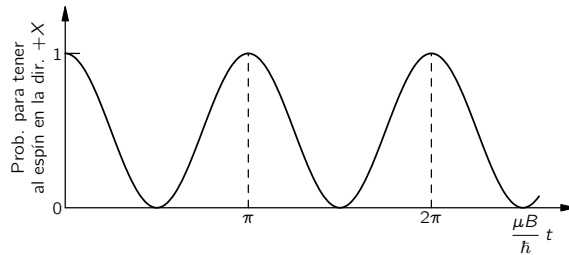
¡Un resultado particularmente simple! Noten que la respuesta concuerda con lo que esperamos para $t = 0$. Obtenemos $A_+(0) = 1$, lo que es correcto, porque supusimos que el muon se encontraba en el estado $(+x)$ en $t = 0$.

La probabilidad P_+ para que el muon se encuentre en el estado $(+x)$ en el tiempo t es $(A_+)^2$ o

$$P_+ = \cos^2 \frac{\mu B}{\hbar} t.$$

La probabilidad oscila entre cero y uno, como se muestra en la figura 7-10. Noten que la probabilidad vuelve a ser uno para $\mu Bt/\hbar = \pi$ (no 2π). Como hemos elevado al cuadrado la función coseno, la probabilidad se repite con la frecuencia $2\mu B/\hbar$.

Figura 7-10. Dependencia temporal de la probabilidad para que una partícula con espín un medio esté en un estado $(+)$ respecto al eje x .



Por lo tanto, encontramos que la probabilidad para atrapar al electrón de la desintegración en el contador de electrones de la figura 7-9, varía periódicamente con la cantidad de tiempo que el muon ha estado dentro del campo magnético. La frecuencia depende del momento magnético μ . El momento magnético del muon ha sido, de hecho, medido de esta manera.

Podemos, naturalmente, usar el mismo método para responder cualquier otra pregunta acerca de la desintegración del muon. Por ejemplo, ¿cómo depende de t la probabilidad para detectar a un electrón de la desintegración en la dirección y a 90° de la dirección x , pero que sigue siendo perpendicular al campo? Si la calculan, la amplitud para estar en el estado $(+y)$ varía como $\cos^2\{(\mu Bt/\hbar) - \pi/4\}$, que oscila con el mismo periodo, pero alcanza su máximo un cuarto de ciclo más tarde, cuando $\mu Bt/\hbar = \pi/4$. En realidad, lo que está sucediendo es que, a medida que pasa el tiempo, el muon atraviesa una sucesión de estados, que corresponden a la polarización completa en una dirección que está girando continuamente alrededor del eje z . Podemos describir esto diciendo que *el espín está en precesión* a la frecuencia

$$\omega_p = \frac{2\mu B}{\hbar}. \quad (7.38)$$

Ya pueden empezar a imaginarse la forma que tomará nuestra descripción cuántica cuando describamos cómo se comportan las cosas en el tiempo.

La matriz hamiltoniana

8-1 Amplitudes y vectores

Antes de iniciar el tema principal de este capítulo, nos gustaría describir algunos conceptos matemáticos muy usados en la literatura de la mecánica cuántica. Al tener una idea sobre ellos, les resultará más fácil leer otros libros o trabajos sobre el tema. La primera idea es el gran parecido matemático entre las ecuaciones de la mecánica cuántica y el producto escalar de dos vectores. Recuerdan que si χ y ϕ son dos estados, la amplitud para iniciar en ϕ y terminar en χ se puede escribir como la suma, sobre un conjunto completo de estados base, de la amplitud para ir de ϕ a uno de los estados base y luego a partir de ese estado base a χ :

$$\langle \chi | \phi \rangle = \sum_{\text{toda } i} \langle \chi | i \rangle \langle i | \phi \rangle. \quad (8.1)$$

Explicamos esto en términos de un aparato de Stern-Gerlach, pero recuerden que no es necesario tener el aparato. La ecuación (8.1) es una ley matemática, es verdadera cuando utilizamos el equipo de filtración, como cuando no lo utilizamos —no siempre es necesario imaginar que el aparato está ahí—. Podemos considerarla simplemente como una fórmula para la amplitud $\langle \chi | \phi \rangle$.

Nos gustaría comparar la ecuación (8.1) con la fórmula del producto escalar de dos vectores $\mathbf{B}$ y $\mathbf{A}$. Si $\mathbf{B}$ y $\mathbf{A}$ son vectores comunes en tres dimensiones, podemos escribir el producto escalar en esta forma:

$$\sum_{\text{toda } i} (\mathbf{B} \cdot \mathbf{e}_i)(\mathbf{e}_i \cdot \mathbf{A}), \quad (8.2)$$

entendiendo que el símbolo $\mathbf{e}_i$ representa los tres vectores unitarios en las direcciones x , y y z . Entonces $\mathbf{B} \cdot \mathbf{e}_1$ es lo que normalmente llamamos B_x ; $\mathbf{B} \cdot \mathbf{e}_2$ es lo que normalmente llamamos B_y ; y así sucesivamente. Por lo tanto, la ecuación (8.2) es equivalente a

$$B_x A_x + B_y A_y + B_z A_z,$$

que es el producto escalar $\mathbf{B} \cdot \mathbf{A}$.

Al comparar las ecuaciones (8.1) y (8.2), podemos ver la siguiente analogía: los estados χ y ϕ corresponden a los dos vectores $\mathbf{A}$ y $\mathbf{B}$. Los estados base i corresponden a los vectores especiales $\mathbf{e}_i$, a los que referimos todos los otros vectores. Cualquier vector se puede representar como combinación lineal de los tres “vectores base” $\mathbf{e}_i$. Aún más, si conocen los coeficientes para cada “vector base” en esta combinación —es decir, sus tres componentes— conocen todo acerca del vector. De una manera análoga, cualquier estado cuántico se puede describir completamente por medio de las amplitudes $\langle i | \phi \rangle$ para ir a los estados base; y si conocen estos coeficientes, saben todo lo que hay que saber acerca del estado. Es por esta estrecha analogía que, a lo llamado “estado”, también se le denomina a menudo “vector de estado”.

Como los vectores base $\mathbf{e}_i$ son perpendiculares, tenemos la relación

$$\mathbf{e}_i \cdot \mathbf{e}_j = \delta_{ij}. \quad (8.3)$$

Esto corresponde a las relaciones (5.25) entre los estados base i ,

$$\langle i | j \rangle = \delta_{ij}. \quad (8.4)$$

Ven ahora por qué se dice que los estados base i son “ortogonales”.

8-1 Amplitudes y vectores

8-2 Descomposición de vectores de estado

8-3 ¿Cuáles son los estados base del mundo?

8-4 ¿Cómo cambian los estados en el tiempo?

8-5 La matriz hamiltoniana

8-6 La molécula de amoniaco

Revisar

Volumen I, capítulo 49,
Modos de vibración

Hay una pequeña diferencia entre la ecuación (8.1) y el producto escalar. Tenemos que

$$\langle \phi | \chi \rangle = \langle \chi | \phi \rangle^*. \quad (8.5)$$

Pero en el álgebra vectorial,

$$\mathbf{A} \cdot \mathbf{B} = \mathbf{B} \cdot \mathbf{A}.$$

Con los números complejos de la mecánica cuántica, tenemos que mantener rigurosamente el orden de los términos, mientras que en el producto escalar el orden no importa.

Consideremos ahora la siguiente ecuación vectorial:

$$\mathbf{A} = \sum_i \mathbf{e}_i (\mathbf{e}_i \cdot \mathbf{A}). \quad (8.6)$$

Es un poco inusual, pero correcta. Significa lo mismo que

$$\mathbf{A} = \sum_i A_i \mathbf{e}_i = A_x \mathbf{e}_x + A_y \mathbf{e}_y + A_z \mathbf{e}_z. \quad (8.7)$$

Observen, no obstante, que la ecuación (8.6) contiene una cantidad que es *diferente* de un producto escalar. El producto escalar es sólo un *número*, mientras que la ecuación (8.6) es una ecuación *vectorial*. Uno de los grandes artificios del álgebra vectorial fue extraer de las ecuaciones el concepto mismo de *vector*. Similarmente, uno estaría tentado a extraer algo análogo a un “vector” de la fórmula cuántica (8.1) —y realmente se puede—. Eliminamos los $\langle \chi |$ de ambos lados de la ecuación (8.1), y escribimos la siguiente ecuación (no se asusten —sólo es notación— y en unos instantes descubrirán lo que significan los símbolos):

$$| \phi \rangle = \sum_i | i \rangle \langle i | \phi \rangle. \quad (8.8)$$

Se considera que el corchete $\langle \chi | \phi \rangle$ está dividido en dos partes. La segunda parte $| \phi \rangle$ se denomina comúnmente *ket*, y la primera parte $\langle \chi |$ se denomina *bra* (juntos hacen un “bra-ket” —una notación propuesta por Dirac); los medios símbolos $| \phi \rangle$ y $\langle \chi |$ son llamados *vectores de estado*. En todo caso, *no* son números y, en general, queremos que los resultados de nuestros cálculos sean números; por consiguiente, esas cantidades “sin terminar” son pasos intermedios en nuestros cálculos.

Hasta el momento hemos escrito todos nuestros resultados en términos de números. ¿Cómo nos las hemos arreglado para evitar los vectores? Es divertido observar que, aun en el álgebra vectorial ordinaria, *podríamos* hacer que en todas las ecuaciones sólo hubiera números. Por ejemplo, en vez de una ecuación vectorial como:

$$\mathbf{F} = m\mathbf{a},$$

siempre podríamos escribir

$$\mathbf{C} \cdot \mathbf{F} = \mathbf{C} \cdot (m\mathbf{a}).$$

Tenemos entonces una ecuación entre productos escalares, la cual es válida para *cualquier* vector $\mathbf{C}$. Pero si es válida para cualquier $\mathbf{C}$, ¿no tiene ningún sentido continuar escribiendo la $\mathbf{C}$!

Volvamos ahora a la ecuación (8.1). Es una ecuación válida para *cualquier* χ . Entonces, para evitar escribir de más, podríamos *quitar* la χ y escribir la ecuación (8.8) en su lugar. Contiene la misma información *siempre* que sobreentendamos que tenemos que “terminarla”, al “multiplicar en la izquierda por” —lo que sólo significa reincorporar— algún $\langle \chi |$ en ambos lados. Así, la ecuación (8.8) significa exactamente lo mismo que la ecuación (8.1) —ni más ni menos—. Cuando quieran números, insertan el $\langle \chi |$ que quieran.

Posiblemente ya se hayan preguntado sobre la ϕ en la ecuación (8.8). Como la ecuación es válida para *cualquier* ϕ , ¿por qué la conservamos? Precisamente, Dirac sugiere que también se extraiga la ϕ , por lo que nos queda solamente

$$| = \sum_i |i\rangle\langle i|. \quad (8.9)$$

¡Y ésta es la gran ley de la mecánica cuántica! (No hay analogías en el análisis vectorial.) Establece que si se *introducen* dos estados χ y ϕ cualesquiera a la izquierda y a la derecha de ambos lados, *obtienen nuevamente* la ecuación (8.1). Realmente no es muy útil, pero es una buena manera de recordar que la ecuación es válida para dos estados cualesquiera.

8-2 Descomposición de vectores de estado

Volvamos a la ecuación (8.8); podemos considerarla de la siguiente manera. Cualquier vector de estado $|\phi\rangle$ se puede representar como una combinación lineal, con los coeficientes apropiados, de un conjunto de “vectores” base —o, si lo prefieren, como una superposición de “vectores unitarios” en proporciones adecuadas—. Para resaltar que los coeficientes $\langle i|\phi\rangle$ son números (complejos) ordinarios, escribimos

$$\langle i|\phi\rangle = C_i.$$

Luego, la ecuación (8.8) es lo mismo que

$$|\phi\rangle = \sum_i |i\rangle C_i. \quad (8.10)$$

Podemos escribir una ecuación similar para cualquier otro vector de estado, digamos $|\chi\rangle$, por supuesto, con coeficientes diferentes — D_i por ejemplo—. Tenemos entonces

$$|\chi\rangle = \sum_i |i\rangle D_i. \quad (8.11)$$

Los D_i son sencillamente las amplitudes $\langle i|\chi\rangle$.

Supongan que comenzamos extrayendo la ϕ de la ecuación (8.1). Obtenemos

$$\langle\chi| = \sum_i \langle\chi|i\rangle\langle i|. \quad (8.12)$$

Recordando que $\langle\chi|i\rangle = \langle i|\chi\rangle^*$, podemos escribir la ecuación anterior en la forma

$$\langle\chi| = \sum_i D_i^* \langle i|. \quad (8.13)$$

Ahora, lo interesante es que simplemente podemos *multiplicar* las ecuaciones (8.13) y (8.10) para recuperar $\langle\chi|\phi\rangle$. Al hacerlo, debemos tener cuidado con los índices de la suma, porque son completamente distintos en las dos ecuaciones. Escribamos primero la ecuación (8.13) en la forma

$$\langle\chi| = \sum_j D_j^* \langle j|,$$

lo cual no cambia nada. Luego, al juntarla con la ecuación (8.10), tenemos

$$\langle\chi|\phi\rangle = \sum_{ij} D_j^* \langle j|i\rangle C_i. \quad (8.14)$$

Recuerden, no obstante, que $\langle j|i\rangle = \delta_{ij}$, por lo que en la suma sólo nos quedan los términos con $j = i$. Obtenemos

$$\langle\chi|\phi\rangle = \sum_i D_i^* C_i, \quad (8.15)$$

donde, por supuesto, $D_i^* = \langle i | \chi \rangle^* = \langle \chi | i \rangle$, y $C_i = \langle i | \phi \rangle$. Observamos nuevamente la estrecha analogía con el producto escalar

$$B \cdot A = \sum_i B_i A_i.$$

La única diferencia es que aparece el complejo conjugado de D_i . Así, la ecuación (8.15) dice que, si se expanden los vectores de estado $\langle \chi |$ y $| \phi \rangle$ en términos de los vectores base $\langle i |$ o $| i \rangle$, la amplitud para ir de ϕ a χ está dada por el tipo de producto escalar en la ecuación (8.15). Esta ecuación es, por supuesto, simplemente la ecuación (8.1) escrita con símbolos diferentes. Por lo tanto, sólo hemos dado vueltas en círculo para acostumbrarnos a los nuevos símbolos.

Quizá convenga insistir que, mientras los vectores del espacio en tres dimensiones están descritos en términos de *tres* vectores unitarios ortogonales, los vectores base $| i \rangle$ de los estados cuánticos deben recorrer el conjunto completo aplicable al problema en cuestión. Según la situación, dos, o tres, o cinco, o un número infinito de estados base se pueden necesitar.

También hemos hablado de lo que ocurre cuando las partículas atraviesan un aparato. Si comenzamos con las partículas en cierto estado ϕ , luego las pasamos a través de un aparato, y después hacemos una medición para ver si están en el estado χ , el resultado está descrito por la amplitud

$$\langle \chi | A | \phi \rangle. \quad (8.16)$$

Este símbolo no tiene análogo en el álgebra vectorial. (Es más parecido a tensores algebraicos, pero la analogía no es especialmente útil.) Vimos en el capítulo 5, ecuación (5.32), que podemos escribir (8.16) en la forma

$$\langle \chi | A | \phi \rangle = \sum_{ij} \langle \chi | i \rangle \langle i | A | j \rangle \langle j | \phi \rangle. \quad (8.17)$$

Esto es sólo un ejemplo de la regla fundamental de la ecuación (8.9), empleada dos veces.

También encontramos que si un aparato B se pone en serie con A , podemos escribir

$$\langle \chi | BA | \phi \rangle = \sum_{ijk} \langle \chi | i \rangle \langle i | B | j \rangle \langle j | A | k \rangle \langle k | \phi \rangle. \quad (8.18)$$

De nuevo, esto proviene directamente del método de Dirac para escribir la ecuación (8.9) —recuerden que siempre podemos colocar una barra ($|$), que es como un factor 1, entre B y A —.

Además, podemos interpretar la ecuación (8.17) de otro modo. Supongan que la partícula entra al aparato A en el estado ϕ y sale de A en el estado ψ (“psi”). En otras palabras, podemos hacernos la pregunta: ¿podemos encontrar una ψ tal que la amplitud para ir de ψ a χ sea siempre, y en todas partes, idénticamente igual a la amplitud $\langle \chi | A | \phi \rangle$? La respuesta es sí. Remplacemos la ecuación (8.17) por

$$\langle \chi | \psi \rangle = \sum_i \langle \chi | i \rangle \langle i | \psi \rangle. \quad (8.19)$$

Es claro que podemos hacer esto si

$$\langle i | \psi \rangle = \sum_j \langle i | A | j \rangle \langle j | \phi \rangle = \langle i | A | \phi \rangle, \quad (8.20)$$

la cual determina a ψ . “Pero no determina a ψ ”, dirán ustedes; “sólo determina $\langle i | \psi \rangle$ ”. Sin embargo, $\langle i | \psi \rangle$ *sí* determina a ψ , porque si se tienen todos los coeficientes que relacionan a ψ con los estados base i , entonces la definición de ψ es única. En realidad, podemos jugar con la notación y escribir el último término de la ecuación (8.20) en la forma

$$\langle i | \psi \rangle = \sum_j \langle i | j \rangle \langle j | A | \phi \rangle. \quad (8.21)$$

Luego, como esta ecuación es válida para toda i , podemos escribir sólo

$$|\psi\rangle = \sum_j |j\rangle \langle j| A |\phi\rangle. \quad (8.22)$$

Entonces podemos decir: “el estado ψ es lo que obtenemos si comenzamos con ϕ y atravesamos el aparato A ”.

Un ejemplo final de los trucos del procedimiento. Empezamos nuevamente con la ecuación (8.17). Como es válida para cualesquiera χ y ϕ , ¡podemos suprimir ambos! Obtenemos entonces†

$$A = \sum_{ij} |i\rangle \langle i| A |j\rangle \langle j|. \quad (8.23)$$

¿Qué significa esto? Significa, ni más y ni menos, lo que obtendrían si volviesen a introducir ϕ y χ . Como está es una ecuación “abierta” e incompleta. Si multiplicamos “en la derecha” por $|\phi\rangle$, se convierte en

$$A |\phi\rangle = \sum_{ij} |i\rangle \langle i| A |j\rangle \langle j| \phi\rangle, \quad (8.24)$$

que es precisamente la ecuación (8.22) de nuevo. En realidad, podríamos haber suprimido las j en esa ecuación y escribir sólo

$$|\psi\rangle = A |\phi\rangle. \quad (8.25)$$

El símbolo A no es una amplitud ni un vector; es una cosa nueva llamada *operador*. Es algo que “opera sobre” un estado para producir un nuevo estado —la ecuación (8.25) dice que $|\psi\rangle$ es lo que resulta cuando A opera sobre $|\phi\rangle$ —. De nuevo, es una ecuación abierta hasta que se complete con algún bra, como $\langle\chi|$ para dar

$$\langle\chi|\psi\rangle = \langle\chi|A|\phi\rangle. \quad (8.26)$$

Por supuesto, el operador A queda completamente especificado si damos la matriz de las amplitudes $\langle i|A|j\rangle$ —que también se escribe A_{ij} — en términos de cualquier conjunto de vectores base.

Realmente no hemos agregado nada nuevo con toda esta nueva notación matemática. Una de las razones para hacer todo esto, fue la de mostrarles cómo escribir partes de ecuaciones, porque en muchos libros encontrarán a las ecuaciones escritas en forma incompleta, y no hay razón para que se paralicen al encontrarlas. Si lo prefieren, siempre pueden añadir las partes que faltan para hacer una ecuación entre números que tenga una apariencia un poco más familiar.

Además, como verán, la notación “bra” y “ket” es muy conveniente. Entre otras cosas, de ahora en adelante podemos identificar a un estado dando su vector de estado. Cuando queramos referirnos a un estado de momentum $\mathbf{p}$ definido, podemos decir: “el estado $|\mathbf{p}\rangle$ ”. O podemos hablar de un estado arbitrario $|\psi\rangle$. Por consistencia, siempre usaremos el ket, escribiendo $|\psi\rangle$, para identificar a un estado. (Es, por supuesto, una elección arbitraria; no habría inconveniente en elegir el bra, $\langle\psi|$.)

8-3 ¿Cuáles son los estados base del mundo?

Hemos descubierto que se puede representar cualquier estado del mundo como una superposición —una combinación lineal con coeficientes adecuados— de estados base. Ustedes podrían preguntar, antes que nada, ¿cuáles estados base? Bueno, hay muchas posibilidades diferentes. Pueden, por ejemplo, proyectar un espín en la dirección z o en cualquier otra dirección. Hay muchas *representaciones* diferentes, que son análogas a los distintos *sistemas de coordenadas* que se pueden

† Podrían pensar que deberíamos escribir $|A|$ en vez de A . Pero entonces se parecería al símbolo del “valor absoluto de A ”, por lo que habitualmente no se escriben las barras. En general, la barra ($|$) se comporta muy parecido al factor uno.

usar para representar a los vectores ordinarios. Entonces, ¿qué coeficientes? Bueno, eso depende de las circunstancias físicas. Diferentes situaciones físicas corresponden diferentes conjuntos de coeficientes. Lo que importa es conocer el “espacio” en el cual están trabajando —en otras palabras, qué significan físicamente los estados base—. Por ello, lo primero que tienen que saber es, en general, cómo son los estados base. Entonces pueden entender cómo describir una situación en términos de estos estados base.

Nos gustaría adelantarnos y hablar un poco de lo que será la descripción cuántica general de la naturaleza —en términos de los conceptos físicos actuales, sin embargo—. Primero, se decide una representación determinada de los estados base —representaciones diferentes siempre son posibles—. Por ejemplo, para una partícula con espín un medio podemos usar los estados más y menos respecto al eje z . Pero el eje z no tiene nada de especial —pueden elegir cualquier otro eje que les guste—. No obstante, siempre elegimos el eje z por consistencia. Supongan que empezamos con una situación con un electrón. Además de las dos posibilidades para el espín (“hacia arriba” y “hacia abajo” en la dirección z), también está el momentum del electrón. Escogemos un conjunto de estados base, cada uno corresponde a un valor del momentum. ¿Y si el electrón no tiene un momentum definido? Está bien; sólo estamos indicando cuáles son los estados *base*. Si el electrón no tiene un momentum definido, tiene una amplitud para tener un momentum y otra amplitud para tener otro momentum, y así sucesivamente. Y si su espín no está necesariamente hacia arriba, tiene una amplitud para que su espín sea hacia arriba con ese momentum, y una amplitud para que su espín sea hacia abajo con el otro momentum, y así sucesivamente. La descripción completa de un electrón sólo requiere, *hasta donde sabemos*, que los estados base estén descritos por el *momentum* y el *espín*. Por lo que un conjunto aceptable de estados base $|i\rangle$ para un solo electrón, se refiere a diferentes valores del momentum y si el espín está hacia arriba o hacia abajo. Diversas mezclas de amplitudes —o sea, distintas combinaciones de las C — describen situaciones diferentes. Lo que un electrón determinado está haciendo, se describe diciendo la amplitud para que tenga espín hacia arriba o espín hacia abajo y un momentum o el otro —para todos los valores posibles del momentum—. Pueden darse cuenta de lo que involucra una descripción cuántica completa de un solo electrón.

¿Y los sistemas con más de un electrón? Entonces los estados base se complican. Supongan que tenemos dos electrones. Tenemos, antes que nada, cuatro estados posibles respecto al espín: ambos electrones con el espín para arriba, el primero para abajo y el segundo para arriba, el primero para arriba y el segundo para abajo, o ambos para abajo. También hay que especificar que el primer electrón tiene el momentum p_1 y que el segundo tiene p_2 . Los estados base de dos electrones requieren la especificación del momentum de ambos, y dos caracteres de espín. Con siete electrones, hay que especificar siete de cada uno.

Si tenemos un protón y un electrón, debemos especificar la dirección del espín del protón y su momentum, y la dirección del espín del electrón y su momentum. Esto es aproximadamente cierto. *Realmente no sabemos* cuál es la representación correcta para el mundo. Está muy bien comenzar suponiendo que si especifican el estado de espín del electrón y su momentum, y lo mismo para el protón, tendrán los estados base; ¿y qué ocurre con las “entrañas” del protón? Veámoslo así. En un átomo de hidrógeno, que tiene un protón y un electrón, tenemos muchos estados base diferentes para describir —espines hacia arriba y hacia abajo para el protón y el electrón, y los diversos valores posibles para el momentum del protón y del electrón—. Luego hay diferentes combinaciones de amplitudes C_i , las cuales describen conjuntamente el carácter del átomo de hidrógeno en diferentes estados. Pero supongan que consideramos al átomo de hidrógeno completo como una “partícula”. Si no supiéramos que el átomo de hidrógeno está hecho de un protón y un electrón, podríamos haber empezado diciendo: “ah, sé cuáles son los estados base —corresponden a un momentum determinado del átomo de hidrógeno—”. No, porque el átomo de hidrógeno tiene partes internas. Puede, por lo tanto, tener varios estados de energía interna diferente, y para describir la naturaleza real se necesita mucho más detalle.

La pregunta es: ¿el protón tiene partes internas? ¿Tenemos que describir un protón dando todos los estados posibles de los protones, y los mesones, y las partículas extrañas? No lo sabemos. Y aunque suponemos que el electrón es simple, de modo que todo lo que tenemos que decir sobre él es su momentum y su espín, puede que mañana descubramos que el electrón también tiene engranajes y ruedas internos. Esto significaría que nuestra representación es incompleta, o incorrecta, o aproximada —del mismo modo que una representación del átomo de hidrógeno que sólo describe su momentum sería incompleta, porque ignoraría el hecho de que el átomo de hidrógeno se puede excitar internamente—. Si un electrón se pudiera excitar internamente y se convirtiera en algo diferente, como un muon por ejemplo, entonces habría que describirlo no sólo dando los estados de la nueva partícula, sino, posiblemente, en términos de algún mecanismo interno más complicado. El *problema principal en el estudio actual de las partículas fundamentales hoy en día* es descubrir cuáles son las representaciones correctas para describir a la naturaleza. Hoy día, *conjeturamos* que para el electrón es suficiente especificar su momentum y su espín. También conjeturamos que hay un protón idealizado, que tiene sus mesones π , y sus mesones K, y así en general, y que todos tienen que ser especificados. Docenas de partículas —¡es una locura!—. La pregunta de cuál *es* una partícula fundamental y cuál *no es* una partícula fundamental —un tema del que se habla mucho actualmente— es la pregunta de cuál será la *representación* final en la descripción cuántica definitiva del mundo. ¿Será el momentum del electrón lo apropiado para describir la naturaleza? Incluso, ¡es correcto enfocar el asunto completo de esta manera! Esta pregunta siempre debe aparecer en cualquier investigación científica. De todos modos, tenemos un problema —cómo encontrar una representación—. Ignoramos la respuesta. Ni siquiera sabemos si tenemos el “verdadero” problema, pero si lo tenemos, primero debemos tratar de descubrir si una partícula determinada es “fundamental” o no.

En la mecánica cuántica no relativista —cuando las energías no son tan altas, y no perturban los mecanismos internos de las partículas extrañas, etc.— se puede hacer un buen trabajo sin preocuparse por todos estos detalles. Pueden decidir sólo especificar el momentum y los espines de los electrones y de los núcleos; luego todo estará bien. En la mayoría de las reacciones químicas, y otros fenómenos de baja energía, no pasa nada en los núcleos; no se excitan. Más aún, si un átomo de hidrógeno se mueve lentamente y choca con suavidad con otros átomos de hidrógeno —siempre sin excitarse internamente, o radiar, o cualquier otra cosa complicada como ésa, y siempre permaneciendo en el estado base de energía para el movimiento interno— pueden usar la aproximación de considerar al átomo de hidrógeno como un objeto, una partícula, y no preocuparse de que *puede* hacer algo en su interior. Ésta será una buena aproximación, siempre y cuando, la energía cinética en cualquier colisión esté muy por debajo de 10 electronvoltios —la energía necesaria para excitar al átomo de hidrógeno a un estado interno diferente—. A menudo haremos una aproximación en la cual no incluimos la posibilidad de movimiento interno, disminuyendo con ello el número de detalles que tenemos que agregar en nuestros estados base. Naturalmente, entonces omitimos algunos fenómenos que aparecerían (comúnmente) a una energía mayor, pero al hacer esas aproximaciones, simplificamos bastante el análisis de los problemas físicos. Por ejemplo, podemos estudiar la colisión de dos átomos de hidrógeno a bajas energías —o cualquier proceso químico— sin preocuparnos de que los núcleos atómicos puedan ser excitados. En suma, cuando podemos despreciar los efectos de cualquier estado interno excitado de una partícula, podemos elegir un conjunto base consistete en los estados de momentum y de la componente z del momentum angular definidos.

Entonces, uno de los problemas al describir la naturaleza es encontrar una representación adecuada para los estados base. Pero ése es sólo el comienzo. Aún queremos ser capaces de decir qué “ocurre”. Si conocemos la “condición” del mundo en un instante, desearíamos conocer la condición en un momento posterior. Por lo tanto, también tenemos que encontrar las leyes que determinan cómo cambian las cosas en el tiempo. Ahora nos dedicaremos a la segunda parte de la estructura de la mecánica cuántica —cómo cambian los estados en el tiempo—.

8-4 ¿Cómo cambian los estados en el tiempo?

Ya hemos hablado sobre cómo podemos representar una situación en la que hacemos que algo atraviese un aparato. Ahora bien, un “aparato” conveniente y encantador es, simplemente, una espera de algunos minutos; es decir, preparan un estado ϕ , y antes de analizarlo, sólo lo dejan reposar. Quizá lo dejen reposar en un campo eléctrico o magnético determinado —depende de las circunstancias físicas del mundo—. De todos modos, cualesquiera que sean las condiciones, dejan al objeto reposar desde el tiempo t_1 hasta el tiempo t_2 . Supongan que sale del primer aparato en la condición ϕ en t_1 . Y luego pasa por un “aparato”, pero el “aparato” consiste simplemente en esperar hasta t_2 . Durante la espera pueden estar pasando varias cosas —fuerzas externas aplicadas u otras cosas—, de modo que algo está sucediendo. Al final de la espera, la amplitud para encontrar al objeto en un estado χ ya no es exactamente la misma que hubiera sido sin la espera. Como “esperar” es sólo un caso especial de un “aparato”, entonces podemos especificar lo que ocurre, dando una amplitud de la misma forma que en la ecuación (8.17). Como la operación de “esperar” es especialmente importante, la llamaremos U en vez de A , y para especificar los tiempos inicial y final, t_1 y t_2 , escribiremos $U(t_2, t_1)$. La amplitud que queremos es

$$\langle \chi | U(t_2, t_1) | \phi \rangle. \quad (8.27)$$

Como cualquier otra amplitud, se puede representar en cualquier sistema de base al escribir

$$\sum_{ij} \langle \chi | i \rangle \langle i | U(t_2, t_1) | j \rangle \langle j | \phi \rangle. \quad (8.28)$$

Entonces U queda descrita por completo al dar todo el conjunto de amplitudes, la matriz

$$\langle i | U(t_2, t_1) | j \rangle. \quad (8.29)$$

Podemos señalar, además, que la matriz $\langle i | U(t_2, t_1) | j \rangle$ da muchos más detalles de los necesarios. El físico teórico de categoría que trabaja en física de altas energías, considera problemas de la siguiente naturaleza general (porque es la manera en que comúnmente se hacen los experimentos). Él comienza con un par de partículas, como un protón y otro protón, que se aproximan desde el infinito. (En el laboratorio, comúnmente una partícula está inmóvil y la otra proviene de un acelerador que está, a nivel atómico, prácticamente en el infinito.) Se estrellan y salen, digamos dos mesones K , seis mesones π y dos neutrones en ciertas direcciones con cierto momentum. ¿Cuál es la amplitud para que esto suceda? Las matemáticas son más o menos así: el estado ϕ especifica los espines y momentum de las partículas incidentes. El estado χ es la pregunta sobre qué sale. Por ejemplo, con qué amplitud se tiene a los seis mesones yendo en tal y cual dirección, y los dos neutrones saliendo en estas otras direcciones, con sus espines de este modo o del otro. En otras palabras, χ estaría especificado dando el momentum, los espines, etc., de los productos finales. Entonces, la tarea del teórico es calcular la amplitud (8.27). Sin embargo, sólo está interesado en el caso particular en el que t_1 es $-\infty$ y t_2 es $+\infty$. (No hay evidencia experimental de los detalles del proceso, sólo de lo que entra y de lo que sale.) El límite de $U(t_2, t_1)$ cuando $t_1 \rightarrow -\infty$ y $t_2 \rightarrow +\infty$ se denomina S , y lo que él quiere es

$$\langle \chi | S | \phi \rangle.$$

O, empleando la forma (8.28), él calcularía la matriz

$$\langle i | S | j \rangle,$$

que se denomina la *matriz* S . Así que, si ven a un físico teórico meditando y diciendo: “todo lo que tengo que hacer es calcular la matriz S ”, ya sabrán qué es lo que le preocupa.

Cómo analizar —cómo especificar las leyes para— la matriz S es una pregunta interesante. En la mecánica cuántica relativista para altas energías, se hace de

una manera, pero en la mecánica cuántica no relativista se puede hacer de otra manera, que resulta muy conveniente. (Esta otra manera también se puede usar en el caso relativista, pero entonces no es tan conveniente.) Consiste en construir la matriz U para un intervalo de tiempo pequeño —o sea, para t_2 y t_1 muy cercanos—. Si podemos encontrar una sucesión de U para intervalos de tiempo sucesivos, podemos ver cómo se comportan las cosas en función del tiempo. Pueden darse cuenta inmediatamente de que esta manera no es muy recomendable para la relatividad, porque no queremos tener que especificar cómo se ve todo “simultáneamente” en todas partes. Pero no nos preocuparemos por eso —sólo trataremos mecánica no relativista—.

Supongan que consideramos a la matriz U correspondiente a una espera de t_1 a t_3 , que es mayor que t_2 . Es decir, tomemos tres tiempos sucesivos: t_1 menor que t_2 menor que t_3 . Podemos decir que la matriz que va entre t_1 y t_3 es el *producto* sucesivo de lo que ocurre al esperar de t_1 a t_2 y luego de t_2 a t_3 . Es como lo que ocurría cuando teníamos dos aparatos B y A en serie. Entonces podemos escribir, siguiendo la notación de la sección 5-6,

$$U(t_3, t_1) = U(t_3, t_2) \cdot U(t_2, t_1). \quad (8.30)$$

En otras palabras, podemos analizar cualquier intervalo de tiempo, si podemos analizar una sucesión de intervalos cortos de tiempo en medio. Sencillamente, multiplicamos todas las piezas; ésta es la manera de analizar la mecánica cuántica en forma no relativista.

Nuestro problema es, entonces, entender a la matriz $U(t_2, t_1)$ para un intervalo infinitesimal de tiempo —para $t_2 = t_1 + \Delta t$ —. Nos preguntamos lo siguiente: si tenemos un estado ϕ ahora, ¿cómo es ese estado en un tiempo infinitesimal Δt más tarde? Veamos cómo podemos escribirlo. Llamemos al estado en el tiempo t , $|\psi(t)\rangle$ (mostramos la dependencia temporal de ψ , para que quede completamente claro que nos referimos a las condiciones en el tiempo t). Ahora preguntamos: ¿cuál es la condición después del pequeño intervalo de tiempo Δt ? La respuesta es

$$|\psi(t + \Delta t)\rangle = U(t + \Delta t, t) |\psi(t)\rangle. \quad (8.31)$$

Esto significa lo mismo que (8.25), es decir, que la amplitud para encontrar χ en el tiempo $t + \Delta t$, es

$$\langle \chi | \psi(t + \Delta t) \rangle = \langle \chi | U(t + \Delta t, t) | \psi(t) \rangle. \quad (8.32)$$

Como aún no tienen práctica en aspectos tan abstractos, proyectemos nuestras amplitudes en una representación determinada. Si multiplicamos ambos lados de la ecuación (8.31) por $\langle i |$, obtenemos

$$\langle i | \psi(t + \Delta t) \rangle = \langle i | U(t + \Delta t, t) | \psi(t) \rangle. \quad (8.33)$$

También podemos descomponer $|\psi(t)\rangle$ en los estados base y escribir

$$\langle i | \psi(t + \Delta t) \rangle = \sum_j \langle i | U(t + \Delta t, t) | j \rangle \langle j | \psi(t) \rangle. \quad (8.34)$$

Podemos entender (8.34) como sigue. Si hacemos que $C_i(t) = \langle i | \psi(t) \rangle$ represente a la amplitud para estar en el estado base i en el tiempo t , entonces podemos considerar que esta amplitud (¡recuerden que es sólo un *número*!) varía en el tiempo. Cada C_i se vuelve una función de t . Y también tenemos información sobre *cómo* las amplitudes C_i varían en el tiempo. Cada amplitud en $(t + \Delta t)$ es proporcional a *todas las otras* amplitudes en el tiempo t multiplicadas por un conjunto de coeficientes. Llamemos U_{ij} a la matriz U , que quiere decir

$$U_{ij} = \langle i | U | j \rangle.$$

Por lo tanto, podemos escribir la ecuación (8.34) en la forma

$$C_i(t + \Delta t) = \sum_j U_{ij}(t + \Delta t, t) C_j(t). \quad (8.35)$$

Ésta será, entonces, la apariencia de la dinámica cuántica.

Todavía no sabemos mucho sobre U_{ij} , excepto por una cosa. Sabemos que si Δt tiende a cero, nada puede ocurrir —obtendríamos simplemente el estado original—. Por lo que, $U_{ii} \rightarrow 1$ y $U_{ij} \rightarrow 0$, si $i \neq j$. En otras palabras, $U_{ij} \rightarrow \delta_{ij}$ cuando $\Delta t \rightarrow 0$. Además, podemos suponer que, para un Δt pequeño, cada coeficiente U_{ij} tiene que diferir de δ_{ij} en cantidades proporcionales a Δt ; entonces podemos escribir

$$U_{ij} = \delta_{ij} + K_{ij} \Delta t. \quad (8.36)$$

No obstante, se acostumbra sacar el factor $(-i/\hbar)^\dagger$ de los coeficientes K_{ij} , por razones históricas y de otra índole; preferimos escribir

$$U_{ij}(t + \Delta t, t) = \delta_{ij} - \frac{i}{\hbar} H_{ij}(t) \Delta t. \quad (8.37)$$

Por supuesto, es lo mismo que la ecuación (8.36) y, si así lo desean, simplemente define a los coeficientes $H_{ij}(t)$. Los términos H_{ij} son sólo las derivadas respecto a t_2 de los coeficientes $U_{ij}(t_2, t_1)$, evaluadas en $t_2 = t_1 = t$.

Al sustituir esta forma de U en la ecuación (8.35), tenemos

$$C_i(t + \Delta t) = \sum_j \left[\delta_{ij} - \frac{i}{\hbar} H_{ij}(t) \Delta t \right] C_j(t). \quad (8.38)$$

Al sumar el término que contiene a δ_{ij} , obtenemos simplemente $C_i(t)$, que podemos pasar al otro lado de la ecuación. Luego dividimos entre Δt , tenemos lo que se puede reconocer como una derivada

$$\frac{C_i(t + \Delta t) - C_i(t)}{\Delta t} = -\frac{i}{\hbar} \sum_j H_{ij}(t) C_j(t)$$

o

$$i\hbar \frac{dC_i(t)}{dt} = \sum_j H_{ij}(t) C_j(t). \quad (8.39)$$

Ustedes recuerdan que $C_i(t)$ es la amplitud $\langle i | \psi \rangle$ para encontrar al estado ψ en uno de los estados base i (al tiempo t). Por lo que la ecuación (8.39) nos indica cómo varía cada uno de los coeficientes $\langle i | \psi \rangle$ en el tiempo. Esto es lo mismo que decir que la ecuación (8.39) nos dice cómo varía el estado ψ respecto al tiempo, ya que estamos describiendo a ψ en términos de las amplitudes $\langle i | \psi \rangle$. La variación de ψ en el tiempo se describe mediante la matriz H_{ij} , que tiene que incluir, por supuesto, lo que le estamos haciendo al sistema para que cambie. Si conocemos los H_{ij} —que contienen la física de la situación y que, en general, pueden depender del tiempo—, entonces tenemos una descripción completa del comportamiento temporal del sistema. Así, la ecuación (8.39) es la ley cuántica de la dinámica del mundo.

(Debemos agregar que siempre tomaremos un conjunto de estados base que estén fijos y no varíen en el tiempo. Hay quienes usan estados base que también varían. Sin embargo, eso es como usar un sistema de coordenadas que rota en la mecánica, y no queremos vernos envueltos en esas complicaciones.)

8-5 La matriz hamiltoniana

La idea es que, por consiguiente, para describir el mundo cuántico, necesitamos escoger un conjunto de estados base i y escribir las leyes físicas dando la matriz de coeficientes H_{ij} . Así, tenemos todo —podemos responder a cualquier pregunta sobre lo que sucederá—. Entonces, tenemos que aprender cuáles son las reglas para encontrar las H correspondientes a cualquier situación física —lo que corresponde a un campo magnético, o a un campo eléctrico, etc.—. Y ésa es la parte más difícil. Por ejemplo, para las nuevas partículas extrañas no tenemos la menor idea

[†] Aquí tenemos ciertas dificultades con la notación. En el factor $(-i/\hbar)$, la i es la unidad imaginaria $\sqrt{-1}$, ¡no el índice i que se refiere al i -ésimo estado base! Esperamos que no les cause confusión.

de qué H_{ij} usar. En otras palabras, nadie conoce las H_{ij} completas para todo el mundo. (Parte del problema es que es muy difícil descubrir las H_{ij} , cuando nadie desconoce cuáles son los estados base.) Tenemos aproximaciones excelentes para fenómenos no relativistas y para algunos otros casos especiales. En particular, tenemos las formas que se necesitan para los movimientos de los electrones en los átomos —para describir la química—. Pero no conocemos la H completa y verdadera para todo el universo.

Los coeficientes H_{ij} se denominan la *matriz hamiltoniana* o, para abreviar, simplemente *el hamiltoniano*. (Cómo es que Hamilton, quien trabajó durante la década de 1830, tiene su nombre relacionado con una matriz cuántica es un cuento muy largo.) Sería mucho mejor llamarla *matriz de energía*, por razones que se verán cuando trabajemos con ella. Así, el problema es: ¡conozcan su hamiltoniano!

El hamiltoniano tiene una propiedad que se puede deducir directamente, es decir, que

$$H_{ij}^* = H_{ji}. \quad (8.40)$$

Esto proviene de la condición de que la probabilidad total para que el sistema esté en *algún* estado, no cambie. Si comienzan con una partícula —un objeto o el mundo— la siguen teniendo a medida que pasa el tiempo. La probabilidad total de encontrarla en alguna parte es

$$\sum_i |C_i(t)|^2,$$

que no debe variar con el tiempo. Si esto es cierto para cualquier condición inicial ϕ , entonces la ecuación (8.40) también debe ser cierta.

Como primer ejemplo, tomemos una situación en la cual las condiciones físicas no varían con el tiempo; queremos decir las condiciones físicas *externas*, de modo que H sea independiente del tiempo. Nadie está encendiendo y apagando imanes. También, elegimos un sistema para el cual se necesita sólo un estado base para la descripción; es una aproximación que podríamos hacer para un átomo de hidrógeno en reposo, o algo similar. La ecuación (8.39) dice, entonces,

$$i\hbar \frac{dC_1}{dt} = H_{11}C_1. \quad (8.41)$$

Sólo una ecuación —¡eso es todo!—. Y si H_{11} es constante, se puede resolver fácilmente esta ecuación diferencial para obtener

$$C_1 = (\text{const.})e^{-(i/\hbar)H_{11}t}. \quad (8.42)$$

Ésta es la dependencia temporal de un estado con una energía $E = H_{11}$ definida. Ven por qué H_{ij} se debería llamar matriz de energía. Es la generalización de la energía para situaciones más complejas.

Ahora, para comprender un poco mejor el significado de las ecuaciones, consideremos un sistema que tiene dos estados de base. La ecuación (8.39) es

$$\begin{aligned} i\hbar \frac{dC_1}{dt} &= H_{11}C_1 + H_{12}C_2, \\ i\hbar \frac{dC_2}{dt} &= H_{21}C_1 + H_{22}C_2. \end{aligned} \quad (8.43)$$

Si las H son de nuevo independientes del tiempo, pueden resolver fácilmente estas ecuaciones. Les corresponde a ustedes intentarlo por diversión, después las retomaremos para resolverlas. Sí, ustedes pueden resolver las ecuaciones de la mecánica cuántica sin conocer las H , en tanto sean independientes del tiempo.

8-6 La molécula de amoníaco

Ahora les mostraremos cómo se puede usar la ecuación dinámica de la mecánica cuántica para describir una circunstancia física especial. Hemos escogido

un ejemplo interesante, pero sencillo, con el cual, haciendo algunas hipótesis razonables sobre el hamiltoniano, podemos lograr ciertos resultados importantes —e incluso prácticos—. Vamos a considerar una situación que se puede describir por medio de dos estados: la molécula de amoníaco.

La molécula de amoníaco tiene un átomo de nitrógeno y tres de hidrógeno ubicados en un plano por debajo del nitrógeno, de modo que la molécula tiene forma de pirámide, como se ha dibujado en la figura 8-1(a). Ahora bien, esta molécula, como cualquier otra, posee un número infinito de estados. Puede girar alrededor de cualquier eje; se puede mover en cualquier dirección; también puede vibrar internamente, etc. Por lo tanto, no es de ninguna manera un sistema de dos estados. Pero intentaremos la aproximación de que todos los otros estados permanecen fijos, porque no intervienen en lo que nos interesa por el momento. Sólo consideraremos que la molécula está girando alrededor de su eje de simetría (como muestra la figura), que no tiene momentum traslacional, y que está vibrando lo menos posible. Esto especifica todas las condiciones excepto una: *aún quedan las dos posiciones posibles para el átomo de nitrógeno* —el nitrógeno puede estar cualquiera de los lados del plano de los átomos de hidrógeno, como se muestra en la figura 8-1(a) y (b)—. Por consiguiente, estudiaremos a la molécula como si fuera un sistema con dos estados. Queremos decir que hay sólo dos estados por los que nos vamos a preocupar, suponiendo que todo lo demás no cambia. Ustedes ven que, aunque sepamos que está girando con cierto momentum angular alrededor del eje, y que se está moviendo con cierto momentum, y que está vibrando de un modo determinado, hay todavía dos estados posibles. Diremos que la molécula está en el estado $|1\rangle$ cuando el nitrógeno está “arriba”, como en la figura 8-1(a), y que está en el estado $|2\rangle$ cuando el nitrógeno está “abajo”, como en (b). Los estados $|1\rangle$ y $|2\rangle$ serán el conjunto de estados base para nuestro análisis del comportamiento de la molécula de amoníaco. En cualquier momento, el estado real $|\psi\rangle$ de la molécula se puede representar dando $C_1 = \langle 1 | \psi \rangle$, la amplitud para estar en el estado $|1\rangle$, y $C_2 = \langle 2 | \psi \rangle$, la amplitud para estar en el estado $|2\rangle$. Así, usando la ecuación (8.8), podemos escribir el vector de estado $|\psi\rangle$ en la forma

$$|\psi\rangle = |1\rangle\langle 1 | \psi \rangle + |2\rangle\langle 2 | \psi \rangle$$

o

$$|\psi\rangle = |1\rangle C_1 + |2\rangle C_2. \quad (8.44)$$

Figura 8-1. Dos arreglos geométricos equivalentes de la molécula de amoníaco.

Ahora bien, lo interesante es que si se sabe que la molécula está en cierto estado en cierto instante, *no* estará en el mismo estado un poco después. Los dos coeficientes C estarán cambiando con el tiempo según la ecuación (8.43) —que es válida para todo sistema de dos estados—. Supongan, por ejemplo, que han hecho una observación —o han hecho una selección de las moléculas— de modo que *saben* que la molécula está *inicialmente* en el estado $|1\rangle$. Un poco más tarde, hay cierta probabilidad para que se encuentre en el estado $|2\rangle$. Para determinar esta probabilidad, tenemos que resolver la ecuación diferencial que nos dice cómo varían las amplitudes en el tiempo.

El único problema es que no sabemos qué usar para los coeficientes H_{ij} en la ecuación (8.43). No obstante, hay algunas cosas que *podemos* decir. Supongan que una vez que la molécula está en el estado $|1\rangle$, no hay posibilidad de que pase al estado $|2\rangle$ y viceversa. Luego entonces H_{12} y H_{21} serían cero, y la ecuación (8.43) será

$$i\hbar \frac{dC_1}{dt} = H_{11}C_1, \quad i\hbar \frac{dC_2}{dt} = H_{22}C_2.$$

Podemos resolver fácilmente estas dos ecuaciones; obtenemos

$$C_1 = (\text{const})e^{-(i/\hbar)H_{11}t}, \quad C_2 = (\text{const})e^{-(i/\hbar)H_{22}t}. \quad (8.45)$$

Éstas son precisamente las amplitudes correspondientes a estados *estacionarios* con energías $E_1 = H_{11}$ y $E_2 = H_{22}$. Observemos que, sin embargo, para la

molécula de amoniaco, los dos estados $|1\rangle$ y $|2\rangle$ tienen una simetría definida. Si la naturaleza es razonable, los elementos de matriz H_{11} y H_{22} deben ser iguales. Los llamaremos E_0 , porque éstos corresponden a la energía que tendrían los estados si H_{12} y H_{21} fueran cero. Pero las ecuaciones (8.45) no nos dice lo que realmente hace la molécula de amoniaco. Resulta que es posible que el nitrógeno se abra camino a través de los tres hidrógenos y pase al otro lado. Es muy difícil; para llegar a la mitad del camino se necesita mucha energía. ¿Cómo puede pasar si no tiene energía suficiente? Hay una cierta amplitud para que penetre la barrera de energía. En la mecánica cuántica, es posible escurrirse rápidamente por una región energéticamente ilegal. Hay, por lo tanto, una pequeña amplitud para que una molécula, que está inicialmente en el estado $|1\rangle$, pase al estado $|2\rangle$. Los coeficientes H_{12} y H_{21} no son realmente cero. De nuevo, por simetría, ambos deben ser iguales —al menos en magnitud—. En realidad, ya sabemos que, en general, H_{ij} debe ser igual al complejo conjugado de H_{ji} , por lo que sólo pueden diferir en la fase. Resulta, como verán, que no se pierde generalidad si las consideramos iguales. Por conveniencia, en lo que sigue las igualamos a un número negativo; hacemos $H_{12} = H_{21} = -A$. Tenemos entonces el siguiente par de ecuaciones:

$$i\hbar \frac{dC_1}{dt} = E_0 C_1 - A C_2, \quad (8.46)$$

$$i\hbar \frac{dC_2}{dt} = E_0 C_2 - A C_1. \quad (8.47)$$

Estas ecuaciones son bastante sencillas y se las puede resolver de varias maneras. Una manera conveniente es la siguiente. Al sumar ambas obtenemos

$$i\hbar \frac{d}{dt} (C_1 + C_2) = (E_0 - A)(C_1 + C_2),$$

cuya solución es

$$C_1 + C_2 = a e^{-(i/\hbar)(E_0 - A)t}. \quad (8.48)$$

Luego, al restar (8.46) y (8.47), encontramos que

$$i\hbar \frac{d}{dt} (C_1 - C_2) = (E_0 + A)(C_1 - C_2),$$

lo cual da

$$C_1 - C_2 = b e^{-(i/\hbar)(E_0 + A)t}. \quad (8.49)$$

Hemos llamado a y b a las dos constantes de integración; por supuesto, se las debe escoger de modo que den la condición inicial de un problema físico particular. Ahora bien, sumando y restando (8.48) y (8.49), obtenemos C_1 y C_2 :

$$C_1(t) = \frac{a}{2} e^{-(i/\hbar)(E_0 - A)t} + \frac{b}{2} e^{-(i/\hbar)(E_0 + A)t}, \quad (8.50)$$

$$C_2(t) = \frac{a}{2} e^{-(i/\hbar)(E_0 - A)t} - \frac{b}{2} e^{-(i/\hbar)(E_0 + A)t}. \quad (8.51)$$

Son iguales excepto por el signo del segundo término.

Tenemos las soluciones, pero ¿qué significan? (El problema en la mecánica cuántica es que no sólo hay que resolver las ecuaciones, sino que hay que entender lo que significan las soluciones!) En primer lugar, observen que si $b = 0$ ambos términos tienen la misma frecuencia $\omega = (E_0 - A)/\hbar$. Si todo cambia con la misma frecuencia, significa que el sistema está en un estado de energía definida —en este caso, la energía $(E_0 - A)$ —. Entonces, hay un estado estacionario de esta energía, en el cual las dos amplitudes C_1 y C_2 son iguales. Llegamos al resultado de que *la molécula de amoniaco tiene una energía definida* $(E_0 - A)$, si hay amplitudes iguales para que el átomo de nitrógeno esté “arriba” o esté “abajo”.

Hay otro estado estacionario posible si $a = 0$; ambas amplitudes tienen entonces la frecuencia $(E_0 + A)/\hbar$. Por lo tanto, hay otro estado de energía

definida ($E_0 + A$) si las dos amplitudes son iguales, pero de signo opuesto; $C_2 = -C_1$. Éstos son los dos únicos estados de energía definida. En el próximo capítulo, discutiremos con más detalle los estados de la molécula de amoniaco; aquí sólo mencionaremos un par de cosas.

Concluimos que *debido* a que hay cierta probabilidad para que el átomo de nitrógeno se cambie de una posición a la otra, la energía no es sólo E_0 , como esperaríamos, sino que hay *dos* niveles de energía ($E_0 + A$) y ($E_0 - A$). Cada uno de los estados posibles de la molécula, cualquiera sea su energía, se “separa” en dos niveles. Decimos *cada* estado porque, como recuerdan, escogimos un estado particular de rotación, y de energía interna, etc. Para cada situación de ese tipo hay un par de niveles de energía debido al cambio en la molécula.

Ahora hagamos la siguiente pregunta sobre la molécula de amoniaco. Supongan que en $t = 0$, sabemos que una molécula está en el estado $|1\rangle$ o, en otras palabras, que $C_1(0) = 1$ y $C_2(0) = 0$. ¿Cuál es la probabilidad para que la molécula se encuentre en el estado $|2\rangle$ en el tiempo t , o que aún se encuentre en el estado $|1\rangle$ en el tiempo t ? Nuestra condición inicial nos dice los valores que deben tener a y b en las ecuaciones (8.50) y (8.51). Si $t = 0$, tenemos que

$$C_1(0) = \frac{a+b}{2} = 1, \quad C_2(0) = \frac{a-b}{2} = 0.$$

Es evidente que $a = b = 1$. Introduciendo estos valores en las expresiones de $C_1(t)$ y $C_2(t)$ y reordenando algunos términos, tenemos

$$C_1(t) = e^{-(i/\hbar)E_0t} \left(\frac{e^{(i/\hbar)At} + e^{-(i/\hbar)At}}{2} \right),$$

$$C_2(t) = e^{-(i/\hbar)E_0t} \left(\frac{e^{(i/\hbar)At} - e^{-(i/\hbar)At}}{2} \right).$$

Podemos reescribir estas ecuaciones en la forma

$$C_1(t) = e^{-(i/\hbar)E_0t} \cos \frac{At}{\hbar}, \quad (8.52)$$

$$C_2(t) = ie^{-(i/\hbar)E_0t} \sin \frac{At}{\hbar}. \quad (8.53)$$

Las dos amplitudes tienen una magnitud que varía armónicamente en el tiempo.

La probabilidad para que la molécula esté en el estado $|2\rangle$ al tiempo t es el cuadrado de la norma de $C_2(t)$:

$$|C_2(t)|^2 = \sin^2 \frac{At}{\hbar}. \quad (8.54)$$

La probabilidad comienza en cero (como debe de ser), aumenta a uno, y luego oscila entre cero y uno, como muestra la curva para P_2 en la figura 8-2. Por supuesto, la probabilidad para estar en el estado $|1\rangle$, no se queda en uno. Se “cae”

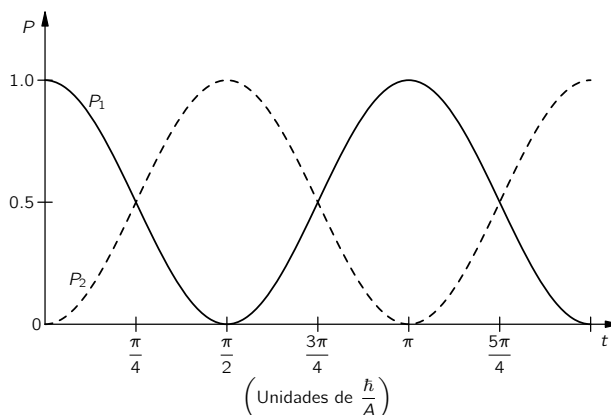


Figura 8-2. La probabilidad P_1 para que una molécula de amoniaco en el estado $|1\rangle$ en $t = 0$, se encuentre en el estado $|1\rangle$ en t . La probabilidad P_2 para que se encuentre en el estado $|2\rangle$.

hacia el segundo estado hasta que la probabilidad para encontrar la molécula en el primer estado es cero, como muestra la curva P_1 de la figura 8-2. La probabilidad va y viene de uno a otro.

Hace mucho vimos lo que ocurre cuando tenemos dos péndulos iguales débilmente acoplados. (Véase capítulo 49, volumen I.) Cuando estiramos uno y lo dejamos ir, oscila, pero luego el otro comienza a oscilar gradualmente. Muy pronto, el segundo péndulo ha tomado toda la energía. Luego, el proceso se invierte, y el péndulo número uno se lleva toda la energía. Ocurre exactamente lo mismo. La rapidez con que se intercambia la energía depende del acoplamiento entre los dos péndulos —la rapidez con que la “oscilación” puede filtrarse de un lado a otro—. También, ustedes recordarán que con los dos péndulos hay dos movimientos especiales —cada uno con una frecuencia definida— que llamamos modos fundamentales. Si estiramos a los dos péndulos juntos, éstos oscilan juntos con una sola frecuencia. Por otro lado, si estiramos uno en un sentido y el otro en el otro, hay otro modo estacionario también de frecuencia definida.

Bueno, aquí tenemos una situación similar —la molécula de amoníaco es matemáticamente como el par de péndulos—. Éstas son las dos frecuencias — $(E_0 - A)/\hbar$ y $(E_0 + A)/\hbar$ — para cuando están oscilando juntas, u oscilando en forma opuesta.

La analogía del péndulo no es mucho más profunda que el principio de que las mismas ecuaciones tienen las mismas soluciones. Las ecuaciones para las amplitudes (8.39) son muy parecidas a las ecuaciones para los osciladores armónicos. (De hecho, ésta es la razón del éxito de nuestra teoría clásica del índice de refracción, en la cual reemplazamos el átomo de la mecánica cuántica por un oscilador armónico, aunque, de manera clásica, ésta no es una imagen razonable de los electrones circulando alrededor del núcleo.) Si tiran del nitrógeno hacia un lado, obtienen una *superposición* de estas dos frecuencias, algo así como un *pulso*, porque el sistema *no* está en uno u otro de los estados de frecuencia definida. La separación de los niveles de energía de la molécula de amoníaco es, sin embargo, un efecto estrictamente cuántico.

La separación de los niveles de energía de la molécula de amoníaco tiene aplicaciones prácticas importantes, de las que nos ocuparemos en el próximo capítulo. ¡Al fin tenemos un ejemplo de un problema físico práctico que pueden comprender con la mecánica cuántica!

El máser† de amoniaco

9-1 Los estados de una molécula de amoniaco

En este capítulo vamos a discutir la aplicación de la mecánica cuántica a un dispositivo práctico, el máser de amoniaco. Pueden preguntarse por qué interrumpimos nuestro desarrollo formal de la mecánica cuántica para resolver un problema especial, pero verán que muchas de las características de este problema especial son muy comunes en la teoría general de la mecánica cuántica, además aprenderán mucho al considerar este problema en detalle. El máser de amoniaco es un dispositivo que genera ondas electromagnéticas, cuyo funcionamiento se basa en las propiedades de la molécula de amoniaco que expusimos brevemente en el capítulo anterior. Comenzamos resumiendo lo que observamos antes.

La molécula de amoniaco tiene muchos estados, pero la estamos considerando como un sistema de dos estados, y sólo pensamos en lo que sucede cuando la molécula está en algún estado específico de rotación o traslación. Un modelo físico para los dos estados se puede visualizar como sigue. Si consideramos que la molécula de amoniaco está girando alrededor de un eje que pasa a través del átomo de nitrógeno, y que es perpendicular al plano de los átomos de hidrógeno, como se muestra en la figura 9-1, aun así hay dos situaciones posibles —el nitrógeno puede estar a un lado del plano de los átomos de hidrógeno o a otro—. Llamamos a estos dos estados $|1\rangle$ y $|2\rangle$. Los tomamos como conjunto de estados base para analizar el comportamiento de la molécula de amoniaco.

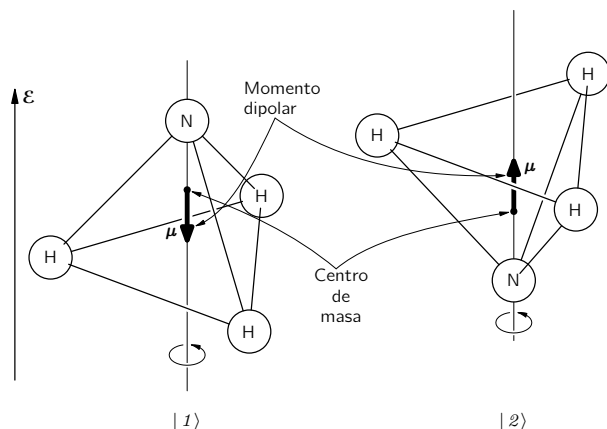


Figura 9-1. Un modelo físico con dos estados base para la molécula de amoniaco. Estos estados tienen el momento dipolar eléctrico μ .

En un sistema con dos estados base, cualquier otro estado $|\psi\rangle$ del sistema siempre se puede expresar como una combinación lineal de los dos estados base; esto es, hay una cierta amplitud C_1 para que esté en un estado base, y una amplitud C_2 de que esté en el otro. Podemos escribir su vector de estado en la forma

$$|\psi\rangle = |1\rangle C_1 + |2\rangle C_2, \quad (9.1)$$

donde

$$C_1 = \langle 1 | \psi \rangle \quad \text{y} \quad C_2 = \langle 2 | \psi \rangle.$$

† N. del T.: La palabra ha sido construida con las iniciales del nombre inglés del fenómeno, amplificación de microondas de una emisión estimulada de radiación.

9-1 Los estados de una molécula de amoniaco

9-2 La molécula en un campo eléctrico estático

9-3 Transiciones en un campo dependiente del tiempo

9-4 Transiciones en resonancia

9-5 Transiciones fuera de resonancia

9-6 Absorción de luz

Estas dos amplitudes cambian con el tiempo de acuerdo a las ecuaciones hamiltonianas (8.43). Al usar la simetría de los dos estados de la molécula de amoníaco, establecemos que $H_{11} = H_{22} = E_0$ y $H_{12} = H_{21} = -A$, y obtenemos como solución [vean las ecuaciones (8.50) y (8.51)]

$$C_1 = \frac{a}{2} e^{-(i/\hbar)(E_0-A)t} + \frac{b}{2} e^{-(i/\hbar)(E_0+A)t}, \quad (9.2)$$

$$C_2 = \frac{a}{2} e^{-(i/\hbar)(E_0-A)t} - \frac{b}{2} e^{-(i/\hbar)(E_0+A)t}. \quad (9.3)$$

Ahora, queremos examinar estas soluciones generales más de cerca. Supongan que la molécula está inicialmente en un estado $|\psi_{II}\rangle$, para el cual el coeficiente b es igual a cero. Por lo tanto, en $t = 0$, las amplitudes para estar en los estados $|1\rangle$ y $|2\rangle$ son idénticas, *y permanecen así para siempre*. Las dos fases varían con el tiempo del mismo modo —con la frecuencia $(E_0 - A)/\hbar$ —. Asimismo, si pusiésemos a la molécula en un estado $|\psi_I\rangle$, para el que $a = 0$, la amplitud C_2 sería el negativo de C_1 , y esta relación permanecería así para siempre. Ambas amplitudes variarían en el tiempo con la frecuencia $(E_0 + A)/\hbar$. Éstas son las dos únicas posibilidades de los estados para los cuales la relación entre el C_1 y C_2 es independiente del tiempo.

Hemos encontrado dos soluciones especiales, en las cuales las dos amplitudes *no varían en magnitud* y, además, tienen fases que varían con la misma frecuencia. Éstos son *estados estacionarios*, según los definimos en la sección 7-1, lo que significa que son *estados con energía definida*. El estado $|\psi_{II}\rangle$ tiene la energía $E_{II} = E_0 - A$, y el estado $|\psi_I\rangle$ la energía $E_I = E_0 + A$. Éstos son los dos únicos estados estacionarios que existen, por lo que encontramos que la molécula tiene dos niveles de energía, con una diferencia de energía $2A$. (Queremos dar a entender, por supuesto, dos niveles de energía para el supuesto estado de rotación y de vibración al que nos referíamos en nuestras suposiciones iniciales.)†

Si no permitiéramos la posibilidad para que el nitrógeno cambie de lugar, hubiésemos tomado A igual a cero, y los dos niveles de energía coincidirían con la energía E_0 . Los niveles reales no son así; su energía *promedio* es E_0 , pero están a una distancia de $\pm A$, con una separación de $2A$ entre las energías de los dos estados. Ya que A es, de hecho, muy pequeña, la diferencia de energía también es muy pequeña.

Para excitar a un *electrón* dentro de un átomo, las energías necesarias son relativamente muy grandes —requiere fotones en el rango visible o ultravioleta—. Para excitar las *vibraciones* de las moléculas, se necesitan fotones en el infrarrojo. Si quieren excitar las *rotaciones*, las diferencias de energías entre los estados corresponden a fotones en el infrarrojo lejano. Pero la diferencia de energía $2A$ es más pequeña que cualquiera de ellas y está, en realidad, por debajo del infrarrojo y dentro de la región de microondas. Experimentalmente, se ha encontrado que hay un par de niveles de energía con una separación de 10^{-4} electronvoltios —que corresponde a una frecuencia de 24 000 megaciclos—. Esto significa, evidentemente, que $2A = hf$, con $f = 24\,000$ megaciclos (que corresponde a una longitud de onda de $1\frac{1}{4}$ cm). Así que tenemos una molécula con una transición que no emite luz en el sentido ordinario, sino que emite microondas.

Para el trabajo que sigue, necesitamos describir estos dos estados de energía definida un poco mejor. Supongan que vamos a construir una amplitud C_{II} tomando la suma de los dos números C_1 y C_2 :

$$C_{II} = C_1 + C_2 = \langle 1 | \Phi \rangle + \langle 2 | \Phi \rangle. \quad (9.4)$$

¿Qué significaría esto? Bueno, es sólo la amplitud para encontrar al estado $|\Phi\rangle$ en un nuevo estado $|\mathit{II}\rangle$, en el cual las amplitudes de los estados base originales

† A continuación es útil —ya sea para ustedes, o para alguien más— tener un modo fácil de distinguir entre los números arábigos 1 y 2 y los romanos I y II. Pensamos que es conveniente reservar los nombres “uno” y “dos” para los números arábigos, y llamar a I y II por los nombres “eins” y “zwei” (¡aunque sería más lógico “unus” y “duo”!).

son iguales. Esto es, escribiendo $C_{II} = \langle II | \Phi \rangle$, podemos extraer el $|\Phi\rangle$ de la ecuación (9.4) —ya que es correcta para cualquier Φ — y obtenemos

$$\langle II | = \langle I | + \langle 2 |,$$

que significa lo mismo que

$$| II \rangle = | I \rangle + | 2 \rangle. \quad (9.5)$$

La amplitud para que el estado $| II \rangle$ esté en el estado $| I \rangle$ es

$$\langle I | II \rangle = \langle I | I \rangle + \langle I | 2 \rangle,$$

que, naturalmente, es sólo 1, ya que $| I \rangle$ y $| 2 \rangle$ son estados base. La amplitud para que el estado $| II \rangle$ esté en el estado $| 2 \rangle$ es también 1, por lo que el estado $| II \rangle$ tiene amplitudes iguales para estar en los dos estados base $| I \rangle$ y $| 2 \rangle$.

Nos encontramos, sin embargo, con un pequeño problema. El estado $| II \rangle$ tiene una probabilidad total más grande que uno para encontrarse en *algún* estado base *o en el otro*. No obstante, esto sólo significa que el vector de estado no está debidamente “normalizado”. Podemos resolverlo recordando que deberíamos tener $\langle II | II \rangle = 1$, y que debe ser así para cualquier estado. Usando la relación general

$$\langle \chi | \Phi \rangle = \sum_i \langle \chi | i \rangle \langle i | \Phi \rangle,$$

haciendo que Φ y χ sean el estado II , y sumando sobre los estados base $| I \rangle$ y $| 2 \rangle$, obtenemos

$$\langle II | II \rangle = \langle II | I \rangle \langle I | II \rangle + \langle II | 2 \rangle \langle 2 | II \rangle.$$

Esto será igual a uno, como debería, si cambiamos nuestra definición de C_{II} —en la ecuación (9.4)— para que sea

$$C_{II} = \frac{1}{\sqrt{2}} [C_1 + C_2].$$

Del mismo modo podemos construir la amplitud

$$C_I = \frac{1}{\sqrt{2}} [C_1 - C_2],$$

o

$$C_I = \frac{1}{\sqrt{2}} [\langle I | \Phi \rangle - \langle 2 | \Phi \rangle]. \quad (9.6)$$

Esta amplitud es la proyección del estado $|\Phi\rangle$ en un nuevo estado $| I \rangle$, que tiene amplitudes opuestas para estar en los estados $| I \rangle$ y $| 2 \rangle$. Es decir, la ecuación (9.6) significa lo mismo que

$$\langle I | = \frac{1}{\sqrt{2}} [\langle I | - \langle 2 |],$$

o

$$| I \rangle = \frac{1}{\sqrt{2}} [| I \rangle - | 2 \rangle], \quad (9.7)$$

de lo cual se sigue que

$$\langle I | I \rangle = \frac{1}{\sqrt{2}} = -\langle 2 | I \rangle.$$

Ahora bien, la razón por la que hemos hecho todo esto, es porque los estados $| I \rangle$ y $| II \rangle$ *se pueden tomar como un nuevo conjunto de estados base*, los cuales son especialmente convenientes para describir los estados estacionarios de la molécula de amoníaco. Recuerdan que el requisito para constituir un conjunto de estados base es

$$\langle i | j \rangle = \delta_{ij}.$$

Ya hemos arreglado las cosas de manera que

$$\langle I | I \rangle = \langle II | II \rangle = 1.$$

Pueden mostrar fácilmente que, de las ecuaciones (9.5) y (9.7), se deduce que

$$\langle I | II \rangle = \langle II | I \rangle = 0.$$

Las amplitudes $C_I = \langle I | \Phi \rangle$ y $C_{II} = \langle II | \Phi \rangle$ para que cualquier estado se encuentre en nuestros nuevos estados base $|I\rangle$ y $|II\rangle$, también deben satisfacer una ecuación hamiltoniana de la forma de la ecuación (8.39). De hecho, si restamos las ecuaciones (9.2) y (9.3) y derivamos respecto a t , vemos que

$$i\hbar \frac{dC_I}{dt} = (E_0 + A)C_I = E_I C_I. \quad (9.8)$$

y sumando las ecuaciones (9.2) y (9.3), tenemos

$$i\hbar \frac{dC_{II}}{dt} = (E_0 - A)C_{II} = E_{II} C_{II}. \quad (9.9)$$

Usando $|I\rangle$ y $|II\rangle$ como estados base, la matriz hamiltoniana adquiere la forma sencilla

$$\begin{aligned} H_{I,I} &= E_I, & H_{I,II} &= 0, \\ H_{II,I} &= 0, & H_{II,II} &= E_{II}. \end{aligned}$$

Noten que cada una de las ecuaciones (9.8) y (9.9) se parece a lo que teníamos en la sección 8-6 para la ecuación de un sistema con un solo estado. Tienen una dependencia temporal simple, que corresponde a una sola energía en forma de exponencial. Cuando transcurre el tiempo, las amplitudes para estar en cada uno de los estados actúan de manera independiente.

Los dos estados estacionarios $|\psi_I\rangle$ y $|\psi_{II}\rangle$ que encontramos antes son, naturalmente, soluciones de las ecuaciones (9.8) y (9.9). El estado $|\psi_I\rangle$ (para el cual $C_1 = -C_2$) tiene

$$C_I = e^{-(i/\hbar)(E_0+A)t}, \quad C_{II} = 0. \quad (9.10)$$

Y el estado $|\psi_{II}\rangle$ (para el cual $C_1 = C_2$) tiene

$$C_I = 0, \quad C_{II} = e^{-(i/\hbar)(E_0-A)t}. \quad (9.11)$$

Recuerden que las amplitudes en la ecuación (9.10) son

$$C_I = \langle I | \psi_I \rangle, \quad \text{y} \quad C_{II} = \langle II | \psi_I \rangle;$$

por lo cual, la ecuación (9.10) significa lo mismo que

$$|\psi_I\rangle = |I\rangle e^{-(i/\hbar)(E_0+A)t}.$$

Es decir, el vector de estado en el estado estacionario $|\psi_I\rangle$ es el mismo que el vector de estado del estado base $|I\rangle$, excepto por el factor exponencial correspondiente a la energía del estado. De hecho para $t = 0$

$$|\psi_I\rangle = |I\rangle;$$

el estado $|I\rangle$ tiene la misma configuración física que el estado estacionario de energía $E_0 + A$. Del mismo modo, tenemos para el segundo estado estacionario que

$$|\psi_{II}\rangle = |II\rangle e^{-(i/\hbar)(E_0-A)t}.$$

El estado $|II\rangle$ es exactamente el estado estacionario de energía $E_0 - A$ para $t = 0$. Por tanto, nuestros dos nuevos estados base $|I\rangle$ y $|II\rangle$ tienen físicamente la forma de los estados de energía definida, sin el factor exponencial del tiempo, para que así puedan ser estados base independientes del tiempo. (En adelante será

conveniente no tener que distinguir siempre entre los estados estacionarios $|\psi_I\rangle$ y $|\psi_{II}\rangle$ y sus estados base $|I\rangle$ y $|II\rangle$, ya que difieren solamente en los factores temporales obvios.)

A manera de resumen, los vectores de estado $|I\rangle$ y $|II\rangle$ son un par de vectores base apropiados para describir los estados de energía definida de la molécula de amoníaco. Se encuentran relacionados con nuestros vectores base originales mediante

$$|I\rangle = \frac{1}{\sqrt{2}}[|1\rangle - |2\rangle], \quad |II\rangle = \frac{1}{\sqrt{2}}[|1\rangle + |2\rangle]. \quad (9.12)$$

Las amplitudes para estar en $|I\rangle$ y $|II\rangle$, se relacionan con C_1 y C_2 mediante

$$C_I = \frac{1}{\sqrt{2}}[C_1 - C_2], \quad C_{II} = \frac{1}{\sqrt{2}}[C_1 + C_2]. \quad (9.13)$$

Cualquier otro estado se puede representar por una combinación lineal de $|1\rangle$ y $|2\rangle$ —con coeficientes C_1 y C_2 — o por una combinación lineal de los estados base de energía definida $|I\rangle$ y $|II\rangle$ —con coeficientes C_I y C_{II} —. Entonces,

$$|\Phi\rangle = |1\rangle C_1 + |2\rangle C_2$$

o

$$|\Phi\rangle = |I\rangle C_I + |II\rangle C_{II}.$$

La segunda forma nos da las amplitudes para encontrar al estado $|\Phi\rangle$ en un estado de energía $E_I = E_0 + A$, o en un estado de energía $E_{II} = E_0 - A$.

9-2 La molécula en un campo eléctrico estático

Si la molécula de amoníaco se encuentra en uno de los dos estados de energía definida, y la perturbamos a una frecuencia ω tal que $\hbar\omega = E_I - E_{II} = 2A$, el sistema puede hacer una transición de un estado a otro. O, si se encuentra en el estado superior, puede pasar al inferior y emitir un fotón. Pero, para inducir estas transiciones, debe haber una conexión física con los estados —algún modo de perturbar el sistema—. Debe haber alguna maquinaria externa que afecte los estados, algo así como un campo magnético o eléctrico. En nuestro caso particular, los estados son sensibles a un campo eléctrico. Por consiguiente, nos enfocaremos en el problema del comportamiento de la molécula de amoníaco en un campo eléctrico externo.

Para discutir el comportamiento en un campo eléctrico, usaremos nuestro sistema de base original $|1\rangle$ y $|2\rangle$, en vez de $|I\rangle$ y $|II\rangle$. Supongan que hay un campo eléctrico en la dirección perpendicular al plano de los átomos de hidrógeno. Al ignorar por el momento la posibilidad de cambios de lugar, ¿sería verdad que la energía de esta molécula es la misma para las dos posiciones del átomo de nitrógeno? En general, no. Los electrones tienden a estar más cerca del nitrógeno que de los núcleos de hidrógeno, por lo que los hidrógenos son ligeramente positivos. La cantidad total depende de los detalles de la distribución electrónica. Es muy complicado saber exactamente cuál es esta distribución, pero, en todo caso, el resultado neto es que la molécula de amoníaco tiene un momento dipolar eléctrico, como se indica en la figura 9-1. Podemos continuar nuestro análisis sin conocer en detalle la dirección o la magnitud del desplazamiento de la carga. Sin embargo, para ser consistentes con la notación de los demás, supongamos que el momento dipolar eléctrico es μ , cuya dirección apunta *desde* el átomo de nitrógeno y es perpendicular al plano de los átomos de hidrógeno.

Ahora bien, cuando el nitrógeno va de un lado al otro, el centro de masa no se moverá, pero el momento dipolar eléctrico cambiará su dirección. Como consecuencia de este momento, la energía en un campo eléctrico $\mathcal{E}$ dependerá de la orientación molecular.† De acuerdo con la suposición que hicimos antes,

† Lamentamos tener que introducir una nueva notación. Como hemos estado usando p y E para el momentum y la energía, no queremos usarlos de nuevo para momento dipolar y campo eléctrico. Recuerden, en esta sección μ es el momento dipolar *eléctrico*.

la energía potencial será mayor si el átomo de nitrógeno apunta en la dirección del campo, y menor si apunta en la dirección contraria; la separación de las dos energías será $2\mu\mathcal{E}$.

En todo lo discutido hasta el momento, hemos supuesto valores de E_0 y de A sin saber cómo calcularlos. Según la teoría física correcta, sería posible calcular estas constantes en términos de las posiciones y los movimientos de todos los núcleos y electrones. Pero nadie lo ha hecho aún. Este sistema consta de 10 electrones y cuatro núcleos, y es un problema demasiado complicado. En realidad, no hay nadie que sepa mucho más acerca de esta molécula que lo que sabemos nosotros. Todo lo que alguien más podría agregar es que, cuando hay un campo eléctrico, la energía de los dos estados es diferente, esta diferencia es proporcional al campo eléctrico. Nosotros hemos llamado al coeficiente de proporcionalidad 2μ , pero su valor se debe determinar experimentalmente. También podemos decir que la molécula tiene una amplitud A para darse vuelta, pero esto se tendría que medir experimentalmente. Nadie puede darnos valores teóricos precisos de μ y de A , porque los cálculos son demasiado complicados como para hacerlos en detalle.

Para la molécula de amoniaco en un campo eléctrico hay que cambiar nuestra descripción. Si ignoramos la amplitud para que la molécula pase de una configuración a la otra, esperaríamos que las energías de los dos estados $|1\rangle$ y $|2\rangle$ fuesen $(E_0 \pm \mu\mathcal{E})$. Siguiendo el procedimiento del capítulo anterior, tomamos

$$H_{11} = E_0 + \mu\mathcal{E}, \quad H_{22} = E_0 - \mu\mathcal{E}. \quad (9.14)$$

También supondremos que el campo eléctrico en cuestión no afecta apreciablemente la geometría de la molécula y que, por lo tanto, no afecta la amplitud para que el nitrógeno pase de una posición a la otra. Podemos decir que H_{12} y H_{21} no cambian; entonces

$$H_{12} = H_{21} = -A. \quad (9.15)$$

Debemos resolver las ecuaciones hamiltonianas (8.43) con estos nuevos valores de H_{ij} . Podríamos resolverlas como antes, pero como nos encontraremos con muchas situaciones en las que queremos las soluciones para sistemas de dos estados, resolvamos las ecuaciones de una vez por todas en el caso general de H_{ij} arbitrarios —suponiendo solamente que no varían en el tiempo—.

Queremos la solución general del par de ecuaciones hamiltonianas

$$i\hbar \frac{dC_1}{dt} = H_{11}C_1 + H_{12}C_2, \quad (9.16)$$

$$i\hbar \frac{dC_2}{dt} = H_{21}C_1 + H_{22}C_2. \quad (9.17)$$

Como son ecuaciones diferenciales lineales con coeficientes constantes, siempre podemos encontrar soluciones que son funciones exponenciales de la variable dependiente t . Primero buscamos una solución en la cual C_1 y C_2 tengan la misma dependencia temporal; podemos tratar con las funciones

$$C_1 = a_1 e^{-i\omega t}, \quad C_2 = a_2 e^{-i\omega t}.$$

Como tal solución corresponde a un estado de energía $E = \hbar\omega$, podemos escribir directamente

$$C_1 = a_1 e^{-(i/\hbar)Et}, \quad (9.18)$$

$$C_2 = a_2 e^{-(i/\hbar)Et}, \quad (9.19)$$

donde E es aún desconocida y hay que determinarla de modo que satisfaga las ecuaciones diferenciales (9.16) y (9.17).

Cuando sustituimos C_1 y C_2 de (9.18) y (9.19) en las ecuaciones diferenciales (9.16) y (9.17), las derivadas nos dan justamente $-iE/\hbar$ por C_1 o C_2 , por lo que los lados izquierdos son sólo EC_1 y EC_2 . Cancelando los dos factores exponenciales comunes, obtenemos

$$Ea_1 = H_{11}a_1 + H_{12}a_2, \quad Ea_2 = H_{21}a_1 + H_{22}a_2.$$

O, reordenando los términos, tenemos

$$(E - H_{11})a_1 - H_{12}a_2 = 0, \quad (9.20)$$

$$-H_{21}a_1 + (E - H_{22})a_2 = 0. \quad (9.21)$$

Para este conjunto de ecuaciones algebraicas homogéneas, habrá soluciones distintas de cero para a_1 y a_2 , solamente si el determinante de los coeficientes de a_1 y a_2 es cero; o sea, si

$$\text{Det} \begin{pmatrix} E - H_{11} & -H_{12} \\ -H_{21} & E - H_{22} \end{pmatrix} = 0. \quad (9.22)$$

Sin embargo, cuando hay sólo dos ecuaciones y dos incógnitas, no necesitamos una idea tan elaborada. Cada una de las dos ecuaciones (9.20) y (9.21) nos da un cociente para los dos coeficientes a_1 y a_2 , y estos dos cocientes deben ser iguales. De la ecuación (9.20) obtenemos

$$\frac{a_1}{a_2} = \frac{H_{12}}{E - H_{11}}, \quad (9.23)$$

y de (9.21) obtenemos

$$\frac{a_1}{a_2} = \frac{E - H_{22}}{H_{21}}. \quad (9.24)$$

Igualando estos dos cocientes, obtenemos que E debe satisfacer

$$(E - H_{11})(E - H_{22}) - H_{12}H_{21} = 0.$$

Éste es el mismo resultado que obtendríamos resolviendo la ecuación (9.22). De cualquier método, tenemos una ecuación cuadrática en E que tiene dos soluciones:

$$E = \frac{H_{11} + H_{22}}{2} \pm \sqrt{\frac{(H_{11} - H_{22})^2}{4} + H_{12}H_{21}}. \quad (9.25)$$

Hay dos valores posibles para la energía E . Noten que ambas soluciones dan *valores reales* para la energía, ya que H_{11} y H_{22} son reales, y $H_{12}H_{21}$ es igual a $H_{12}H_{12}^* = |H_{12}|^2$, que es real y positivo.

Usando la misma convención que antes, llamaremos a la energía superior E_I y a la energía inferior E_{II} . Tenemos

$$E_I = \frac{H_{11} + H_{22}}{2} + \sqrt{\frac{(H_{11} - H_{22})^2}{4} + H_{12}H_{21}}, \quad (9.26)$$

$$E_{II} = \frac{H_{11} + H_{22}}{2} - \sqrt{\frac{(H_{11} - H_{22})^2}{4} + H_{12}H_{21}}. \quad (9.27)$$

Usando separadamente cada una de estas dos energías en las ecuaciones (9.18) y (9.19), tenemos las amplitudes para los dos estados estacionarios (los estados de energía definida). Si no hay perturbaciones externas, un sistema inicialmente en uno de estos estados permanecerá en él para siempre —sólo su fase cambia—.

Podemos comprobar nuestros resultados en dos casos especiales. Si $H_{12} = H_{21} = 0$, tenemos que $E_I = H_{11}$ y $E_{II} = H_{22}$. Esto es correcto, porque entonces las ecuaciones (9.16) y (9.17) no están acopladas, y cada una representa un estado de energía H_{11} y H_{22} . También, si decimos que $H_{11} = H_{22} = E_0$ y $H_{21} = H_{12} = -A$, obtenemos la solución que encontramos antes:

$$E_I = E_0 + A \quad \text{y} \quad E_{II} = E_0 - A.$$

Para el caso general, las dos soluciones E_I y E_{II} se refieren a dos estados, que podemos llamar de nuevo los estados

$$|\psi_I\rangle = |I\rangle e^{-(i/\hbar)E_I t} \quad \text{y} \quad |\psi_{II}\rangle = |II\rangle e^{-(i/\hbar)E_{II} t}.$$

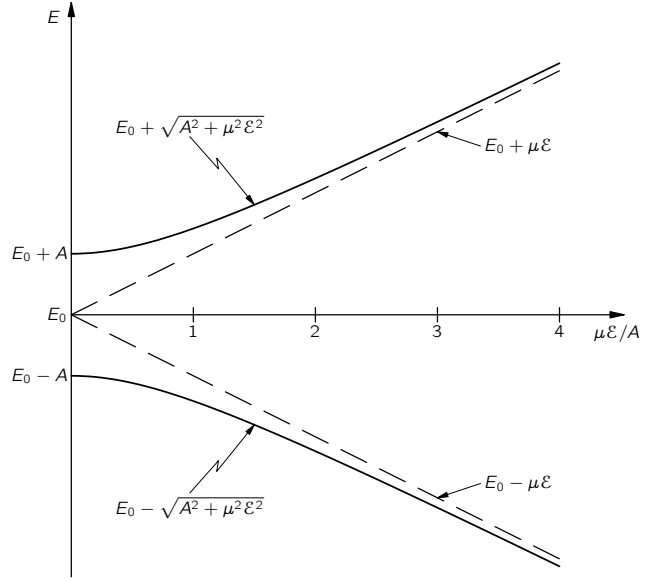


Figura 9-2. Niveles de energía de la molécula de amoníaco en un campo eléctrico.

Estos estados tendrán C_1 y C_2 como en las ecuaciones (9.18) y (9.19), donde aún hay que determinar a_1 y a_2 . Su cociente está dado ya sea por (9.23) o (9.24). También deben satisfacer una condición más. Si se sabe que el sistema está en uno de los estados estacionarios, la suma de las probabilidades de que lo encontremos en $|1\rangle$ o en $|2\rangle$ debe ser igual a uno. Debemos tener que

$$|C_1|^2 + |C_2|^2 = 1, \quad (9.28)$$

o, lo que es equivalente,

$$|a_1|^2 + |a_2|^2 = 1. \quad (9.29)$$

Estas condiciones no especifican a a_1 y a_2 de manera única; aún quedan indeterminadas en una fase arbitraria —en otras palabras, en un factor como $e^{i\delta}$ —. Aunque se puede dar soluciones generales para las a_i ,† es usualmente más conveniente calcularlas para cada caso especial.

Volvamos ahora a nuestro ejemplo de la molécula de amoníaco en un campo eléctrico. Usando los valores para H_{11} , H_{22} y H_{12} dados en (9.14) y (9.15), obtenemos para las energías de los dos estados estacionarios

$$E_I = E_0 + \sqrt{A^2 + \mu^2 \mathcal{E}^2}, \quad E_{II} = E_0 - \sqrt{A^2 + \mu^2 \mathcal{E}^2}. \quad (9.30)$$

Estas dos energías están representadas en función de la intensidad del campo eléctrico $\mathcal{E}$ en la figura 9-2. Por supuesto, cuando el campo eléctrico es cero, las dos energías son simplemente $E_0 \pm A$. Cuando se aplica un campo eléctrico, la separación entre los dos niveles se hace mayor. La separación aumenta primero muy despacio con $\mathcal{E}$, pero eventualmente se hace proporcional a $\mathcal{E}$. (La curva es una hipérbola.) Para campos extremadamente grandes, las energías son simplemente

$$E_I = E_0 + \mu \mathcal{E} = H_{11}, \quad E_{II} = E_0 - \mu \mathcal{E} = H_{22}. \quad (9.31)$$

El hecho de que haya una amplitud para que el nitrógeno cambie de lugar, tiene un efecto muy pequeño cuando las dos posiciones tienen energías muy diferentes. Éste es un punto interesante al que volveremos más adelante.

† Por ejemplo, el conjunto siguiente es una solución aceptable, como pueden verificar fácilmente:

$$a_1 = \frac{H_{12}}{[(E - H_{11})^2 + H_{12}H_{21}]^{1/2}}, \quad a_2 = \frac{E - H_{11}}{[(E - H_{11})^2 + H_{12}H_{21}]^{1/2}}.$$

Finalmente estamos preparados para entender el funcionamiento del máser de amoniaco. La idea es la siguiente. Primero, encontramos un modo de separar las moléculas en el estado $|I\rangle$ de aquellas en el estado $|II\rangle$.[†] Después, las moléculas en el estado de mayor energía $|I\rangle$ se hacen pasar a través de una cavidad que tiene una frecuencia de resonancia de 24 000 megaciclos. Las moléculas pueden liberar energía en la cavidad —de una manera que discutiremos más adelante— y salir de la cavidad en el estado $|II\rangle$. Cada molécula que haga esta transición libera la energía $E = E_I - E_{II}$ en la cavidad. Además, la energía de las moléculas aparecerá como energía eléctrica en la cavidad.

¿Cómo podemos separar los dos estados moleculares? Un método es el siguiente. El gas de amoniaco sale de un pequeño inyector y atraviesa un par de rejillas para producir un haz estrecho, como se muestra en la figura 9-3. Después, el haz atraviesa una región en la que hay un gran campo eléctrico transversal. Los electrodos que producen el campo tienen una forma tal que el campo eléctrico varía rápidamente a través del haz. Así, el cuadrado del campo eléctrico $\mathcal{E} \cdot \mathcal{E}$ tendrá un gradiente grande y perpendicular al haz. De manera que una molécula en el estado $|I\rangle$ tiene una energía que aumenta con $\mathcal{E}^2$, por lo que esta parte del haz se desviará hacia la región de menor $\mathcal{E}^2$. Por otra parte, una molécula en el estado $|II\rangle$ será desviada hacia la región de mayor $\mathcal{E}^2$, ya que su energía disminuye cuando $\mathcal{E}^2$ aumenta.

Además, con los campos eléctricos que se puede tener en el laboratorio, la energía $\mu\mathcal{E}$ es siempre mucho más pequeña que A . En estos casos, la raíz cuadrada en la ecuación (9.30) se puede aproximar por

$$A \left(1 + \frac{1}{2} \frac{\mu^2 \mathcal{E}^2}{A^2} \right). \quad (9.32)$$

Por lo que los niveles de energía, para fines prácticos, son

$$E_I = E_0 + A + \frac{\mu^2 \mathcal{E}^2}{2A} \quad (9.33)$$

y

$$E_{II} = E_0 - A - \frac{\mu^2 \mathcal{E}^2}{2A}. \quad (9.34)$$

Y las energías varían en una forma aproximadamente lineal con $\mathcal{E}^2$. La fuerza sobre las moléculas es entonces

$$\mathbf{F} = \frac{\mu^2}{2A} \nabla \mathcal{E}^2. \quad (9.35)$$

Muchas moléculas tienen una energía en un campo eléctrico, la cual es proporcional a $\mathcal{E}^2$. El coeficiente es la polarizabilidad de la molécula. El amoniaco tiene una polarizabilidad inusualmente alta debido al pequeño valor de A en el denominador. Por eso, las moléculas de amoniaco son extremadamente sensibles a un campo eléctrico. (¿Qué esperarían para el coeficiente dieléctrico del gas NH_3 ?)

9-3 Transiciones en un campo dependiente del tiempo

En el máser de amoniaco, el haz con moléculas en el estado $|I\rangle$ y con energía E_I se hace pasar a través de una cavidad resonante, como se muestra en la figura 9-4. Descartamos al otro haz. Dentro de la cavidad, hay un campo eléctrico que varía con el tiempo, por lo que el próximo problema a discutir es el comportamiento de una molécula en un campo eléctrico que varía con el tiempo. Tenemos un problema completamente diferente —uno con un hamiltoniano que varía en el tiempo—. Ya que H_{ij} depende de $\mathcal{E}$, las H_{ij} varían con el tiempo, y debemos determinar el comportamiento del sistema en estas circunstancias.

[†] De ahora en adelante escribiremos $|I\rangle$ y $|II\rangle$ en lugar de $|\psi_I\rangle$ y $|\psi_{II}\rangle$. Deben recordar que los estados verdaderos $|\psi_I\rangle$ y $|\psi_{II}\rangle$ son los estados base de energía multiplicados por el factor exponencial apropiado.

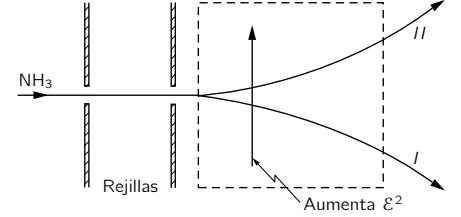


Figura 9-3. El haz de amoniaco se puede separar mediante un campo eléctrico, en el que $\mathcal{E}^2$ tiene un gradiente perpendicular al haz.

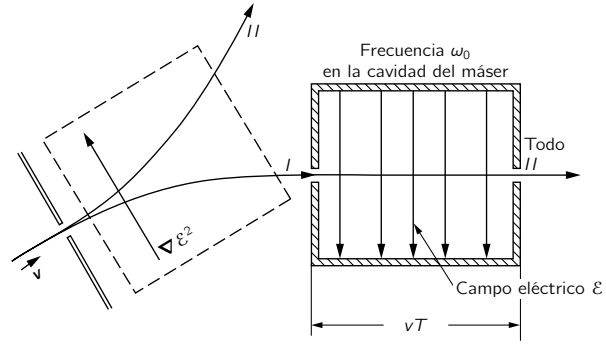


Figura 9-4. Diagrama esquemático del máser de amoníaco.

Para comenzar, escribimos las ecuaciones que tenemos que resolver:

$$\begin{aligned} i\hbar \frac{dC_1}{dt} &= (E_0 + \mu\mathcal{E})C_1 - AC_2, \\ i\hbar \frac{dC_2}{dt} &= -AC_1 + (E_0 - \mu\mathcal{E})C_2. \end{aligned} \quad (9.36)$$

Para ser precisos, supongamos que el campo eléctrico varía de manera sinusoidal; así podemos escribir

$$\mathcal{E} = 2\mathcal{E}_0 \cos \omega t = \mathcal{E}_0(e^{i\omega t} + e^{-i\omega t}). \quad (9.37)$$

En funcionamiento, la frecuencia ω será aproximadamente igual a la frecuencia de resonancia de la transición molecular $\omega_0 = 2A/\hbar$, pero por ahora lo trataremos en una forma general, así que puede tener cualquier valor. La mejor manera para resolver nuestras ecuaciones es formando combinaciones lineales de C_1 y C_2 , como hicimos antes. Así que, sumamos las dos ecuaciones, las dividimos entre la raíz cuadrada de 2, y usamos las definiciones de C_I y C_{II} que teníamos en la ecuación (9.13). Obtenemos

$$i\hbar \frac{dC_{II}}{dt} = (E_0 - A)C_{II} + \mu\mathcal{E}C_I. \quad (9.38)$$

Notarán que esta ecuación es igual a la ecuación (9.9), con un término adicional debido al campo eléctrico. Del mismo modo, si restamos las dos ecuaciones de (9.36), obtenemos

$$i\hbar \frac{dC_I}{dt} = (E_0 + A)C_I + \mu\mathcal{E}C_{II}. \quad (9.39)$$

Ahora la pregunta es, ¿cómo resolver estas ecuaciones? Son más difíciles que nuestro conjunto anterior, ya que $\mathcal{E}$ depende de t ; y de hecho, para un $\mathcal{E}(t)$ general, la solución no se puede expresar mediante funciones elementales. Sin embargo, podemos obtener una buena aproximación siempre que el campo eléctrico sea pequeño. En primer lugar escribiremos

$$\begin{aligned} C_I &= \gamma_I e^{-i(E_0+A)t/\hbar} = \gamma_I e^{-i(E_I)t/\hbar}, \\ C_{II} &= \gamma_{II} e^{-i(E_0-A)t/\hbar} = \gamma_{II} e^{-i(E_{II})t/\hbar}. \end{aligned} \quad (9.40)$$

Si no hubiese campo eléctrico, estas soluciones serían correctas sólo escogiendo γ_I y γ_{II} como dos constantes complejas. De hecho, ya que la probabilidad para estar en el estado $|I\rangle$ es el cuadrado de la norma de C_I y para estar en el estado $|II\rangle$ es el cuadrado de la norma de C_{II} , entonces la probabilidad para estar en el estado $|I\rangle$ o en el estado $|II\rangle$ es sólo $|\gamma_I|^2$ y $|\gamma_{II}|^2$. Por ejemplo, si el sistema empezara originalmente en el estado $|II\rangle$, de tal modo que γ_I fuese cero y $|\gamma_{II}|^2$ fuese uno, este estado continuaría así para siempre. No habría posibilidad de que la molécula, originalmente en el estado $|II\rangle$, pasase alguna vez al estado $|I\rangle$.

La idea de escribir nuestras ecuaciones en la forma (9.40) es que, si $\mu\mathcal{E}$ es pequeño comparado con A , las soluciones aún se pueden escribir de esta forma, pero entonces γ_I y γ_{II} se hacen funciones del tiempo que varían lentamente—donde por “varían lentamente” queremos decir lentamente *en comparación* con las funciones exponenciales—. Éste es el truco. Usamos el hecho de que γ_I y γ_{II} varían lentamente para obtener una solución aproximada.

Ahora queremos sustituir C_I de (9.40) en la ecuación diferencial (9.39), pero debemos recordar que γ_I es también una función de t . Tenemos

$$i\hbar \frac{dC_I}{dt} = E_I \gamma_I e^{-iE_I t/\hbar} + i\hbar \frac{d\gamma_I}{dt} e^{-iE_I t/\hbar}.$$

La ecuación diferencial se convierte en

$$\left(E_I \gamma_I + i\hbar \frac{d\gamma_I}{dt} \right) e^{-(i/\hbar)E_I t} = E_I \gamma_I e^{-(i/\hbar)E_I t} + \mu\mathcal{E} \gamma_{II} e^{-(i/\hbar)E_{II} t}. \quad (9.41)$$

Del mismo modo, la ecuación en dC_{II}/dt se convierte en

$$\left(E_{II} \gamma_{II} + i\hbar \frac{d\gamma_{II}}{dt} \right) e^{-(i/\hbar)E_{II} t} = E_{II} \gamma_{II} e^{-(i/\hbar)E_{II} t} + \mu\mathcal{E} \gamma_I e^{-(i/\hbar)E_I t}. \quad (9.42)$$

Ahora notarán que tenemos términos iguales en los dos lados de cada ecuación. Cancelamos estos términos, y multiplicamos la primera ecuación por $e^{+iE_I t/\hbar}$ y la segunda por $e^{+iE_{II} t/\hbar}$. Recordando que $(E_I - E_{II}) = 2A = \hbar\omega_0$, obtenemos finalmente,

$$\begin{aligned} i\hbar \frac{d\gamma_I}{dt} &= \mu\mathcal{E}(t) e^{i\omega_0 t} \gamma_{II}, \\ i\hbar \frac{d\gamma_{II}}{dt} &= \mu\mathcal{E}(t) e^{-i\omega_0 t} \gamma_I. \end{aligned} \quad (9.43)$$

Ahora tenemos un par de ecuaciones aparentemente simples—y, naturalmente, todavía son exactas—. La derivada de una variable es una función del tiempo $\mu\mathcal{E}(t)e^{i\omega_0 t}$, multiplicada por la segunda variable; la derivada de la segunda es una función semejante del tiempo, multiplicada por la primera. Aunque estas ecuaciones simples no se pueden resolver en general, nosotros lo haremos para algunos casos especiales.

Estamos, al menos por el momento, interesados solamente en el caso de un campo eléctrico oscilante. Tomando $\mathcal{E}(t)$ como está dado en la ecuación (9.37), encontramos que las ecuaciones para γ_I y γ_{II} se convierten en

$$\begin{aligned} i\hbar \frac{d\gamma_I}{dt} &= \mu\mathcal{E}_0 [e^{i(\omega+\omega_0)t} + e^{-i(\omega-\omega_0)t}] \gamma_{II}, \\ i\hbar \frac{d\gamma_{II}}{dt} &= \mu\mathcal{E}_0 [e^{i(\omega-\omega_0)t} + e^{-i(\omega+\omega_0)t}] \gamma_I. \end{aligned} \quad (9.44)$$

Ahora bien, si $\mathcal{E}_0$ es suficientemente pequeño, las rapidezces de variación de γ_I y γ_{II} también son pequeñas. Las dos γ no variarán mucho con t , especialmente en comparación con las variaciones rápidas debidas a los términos exponenciales. Estos términos exponenciales tienen partes reales e imaginarias que oscilan con la frecuencia $\omega + \omega_0$ o $\omega - \omega_0$. Los términos con $\omega + \omega_0$ oscilan muy rápido alrededor del valor promedio de cero y, por lo tanto, no contribuyen mucho en promedio a la rapidez de variación de γ . Así que podemos hacer una aproximación razonablemente buena reemplazando estos términos por su valor promedio, o sea, cero. Sólo no los tomamos en cuenta, y usamos como aproximación:

$$\begin{aligned} i\hbar \frac{d\gamma_I}{dt} &= \mu\mathcal{E}_0 e^{-i(\omega-\omega_0)t} \gamma_{II}, \\ i\hbar \frac{d\gamma_{II}}{dt} &= \mu\mathcal{E}_0 e^{i(\omega-\omega_0)t} \gamma_I. \end{aligned} \quad (9.45)$$

Aún los términos que nos quedan, con exponentes proporcionales a $(\omega - \omega_0)$, variarán también rápidamente a menos que ω sea cercana a ω_0 . Solamente así, los

lados derechos variarán suficientemente lento como para que se acumule cualquier cantidad apreciable cuando integremos las ecuaciones respecto a t . En otras palabras, con un campo eléctrico *débil*, las únicas frecuencias significativas son las cercanas a ω_0 .

Con la aproximación hecha para llegar a (9.45), las ecuaciones se pueden resolver de manera exacta, pero el procedimiento es un poco laborioso, de modo que no lo haremos sino hasta más adelante, cuando tratemos otro problema del mismo tipo. Ahora sólo las resolveremos aproximadamente —o mejor dicho, encontraremos una solución exacta para el caso de resonancia perfecta, $\omega = \omega_0$, y una solución aproximada para frecuencias próximas a la resonancia—.

9-4 Transiciones en resonancia

Primero consideremos el caso de resonancia perfecta. Si hacemos $\omega = \omega_0$, los exponenciales son iguales a uno en las dos ecuaciones (9.45), y sólo nos queda

$$\frac{d\gamma_I}{dt} = -\frac{i\mu\mathcal{E}_0}{\hbar} \gamma_{II}, \quad \frac{d\gamma_{II}}{dt} = -\frac{i\mu\mathcal{E}_0}{\hbar} \gamma_I. \quad (9.46)$$

Si eliminamos primero a γ_I y luego a γ_{II} de estas ecuaciones, encontramos que cada una de ellas satisface la ecuación diferencial del movimiento armónico simple:

$$\frac{d^2\gamma}{dt^2} = -\left(\frac{\mu\mathcal{E}_0}{\hbar}\right)^2 \gamma. \quad (9.47)$$

Las soluciones generales de estas ecuaciones se pueden formar con senos y cosenos. Como pueden verificar fácilmente, las siguientes ecuaciones son una solución:

$$\begin{aligned} \gamma_I &= a \cos\left(\frac{\mu\mathcal{E}_0}{\hbar}t\right) + b \sin\left(\frac{\mu\mathcal{E}_0}{\hbar}t\right), \\ \gamma_{II} &= ib \cos\left(\frac{\mu\mathcal{E}_0}{\hbar}t\right) - ia \sin\left(\frac{\mu\mathcal{E}_0}{\hbar}t\right), \end{aligned} \quad (9.48)$$

donde a y b son constantes que hay que determinar para que se ajusten a cualquier situación física en particular.

Por ejemplo, supongan que a $t = 0$ nuestro sistema molecular se encontraba en el estado superior de energía $|I\rangle$, lo cual requeriría —según la ecuación (9.40)— que $\gamma_I = 1$ y $\gamma_{II} = 0$ en $t = 0$. Para esta situación necesitaríamos que $a = 1$ y $b = 0$. La probabilidad para que la molécula esté en el estado $|I\rangle$ en un t posterior es el cuadrado de la norma de γ_I , o sea

$$P_I = |\gamma_I|^2 = \cos^2\left(\frac{\mu\mathcal{E}_0}{\hbar}t\right). \quad (9.49)$$

Del mismo modo, la probabilidad para que la molécula esté en el estado $|II\rangle$ está dada por el cuadrado de la norma de γ_{II} ,

$$P_{II} = |\gamma_{II}|^2 = \sin^2\left(\frac{\mu\mathcal{E}_0}{\hbar}t\right). \quad (9.50)$$

Siempre que $\mathcal{E}$ sea pequeño y estemos en resonancia, las probabilidades están dadas por funciones oscilatorias sencillas. La probabilidad para estar en el estado $|I\rangle$ cae desde uno hasta cero y vuelve a uno, mientras que la probabilidad para estar en el estado $|II\rangle$ sube desde cero hasta uno y vuelve a caer. Las variaciones temporales de las dos probabilidades se muestran en la figura 9-5. No hay necesidad de decir, la suma de las dos probabilidades es siempre igual a uno; ¡la molécula está siempre en *algún* estado!

Supongamos que la molécula tarda un tiempo T para atravesar la cavidad. Si hacemos la cavidad lo suficientemente larga como para que $\mu\mathcal{E}_0 T/\hbar = \pi/2$, entonces una molécula que entra en el estado $|I\rangle$, saldrá en el estado $|II\rangle$. Si entra en la cavidad en el estado superior, saldrá de ella en el estado inferior.

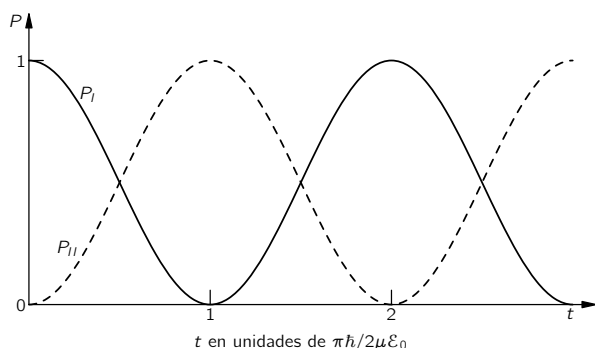


Figura 9-5. Probabilidades para los dos estados de la molécula de amoníaco en un campo eléctrico sinusoidal.

En otras palabras, su energía ha disminuido, y la energía perdida no puede ir a ningún otro lugar que no sea al mecanismo que genera el campo. Los detalles de cómo la energía de la molécula se introduce en las oscilaciones de la cavidad no son simples; sin embargo, no necesitamos estudiar estos detalles, ya que podemos usar el principio de conservación de la energía. (Podríamos estudiarlos si tuviésemos que hacerlo, pero entonces tendríamos que tratar con la mecánica cuántica del campo en la cavidad, además de la mecánica cuántica del átomo.)

En resumen: la molécula entra en la cavidad, al campo de la cavidad —que oscila exactamente a la frecuencia apropiada— induce transiciones del estado superior al inferior, y la energía liberada se introduce en el campo oscilante. En un máser en funcionamiento, las moléculas liberan energía suficiente para mantener las oscilaciones en la cavidad —no sólo proveen la potencia suficiente para suplir las pérdidas de la cavidad, sino también provee pequeñas cantidades de potencia en exceso que se puede sacar de la cavidad—. Esto es, la energía molecular se convierte en energía de un campo electromagnético externo.

Recuerden que, antes de que el haz entre en la cavidad, tenemos que usar un filtro para separar el haz de modo que sólo entre el estado superior. Es fácil demostrar que si ustedes comienzan con moléculas en el estado inferior, el proceso se realiza en la otra dirección y se toma energía de la cavidad. Si se hace pasar un haz sin filtrar, hay tantas moléculas liberando energía como tomándola, así que no sucedería gran cosa. En funcionamiento real no es necesario, por supuesto, hacer a $(\mu\mathcal{E}_0T/\hbar)$ exactamente igual a $\pi/2$. Para cualquier otro valor (excepto para un múltiplo entero exacto de π), hay cierta probabilidad para transiciones del estado $|I\rangle$ al $|II\rangle$. Para otros valores, sin embargo, el dispositivo no es 100 % eficiente; muchas de las moléculas que salen de la cavidad podrían haber liberado cierta energía, pero no lo hicieron.

En usos reales, la velocidad de todas las moléculas no es la misma; tienen una especie de distribución de Maxwell. Esto significa que los periodos ideales para moléculas diferentes serán distintos, y es imposible obtener 100 % de eficiencia para todas las moléculas al mismo tiempo. Además, hay otra complicación que es fácil de tener en cuenta, pero no queremos preocuparnos con ella en esta etapa. Recuerdan que el campo eléctrico en una cavidad por lo general varía de punto a punto a lo largo de la cavidad. Por eso, cuando las moléculas avanzan por la cavidad, el campo eléctrico en la molécula varía en una forma más complicada que la oscilación sinusoidal sencilla que supusimos. Es claro que tendríamos que usar una integración más complicada para hacer el problema de manera exacta, pero aun así la idea general es la misma.

Hay otras formas de hacer máseres. En lugar de separar los átomos en el estado $|I\rangle$ de aquellos en el estado $|II\rangle$ mediante un aparato de Stern-Gerlach, se podría tener a los átomos ya en la cavidad (en forma de gas o de sólido) y cambiar a los átomos del estado $|II\rangle$ al estado $|I\rangle$ de alguna manera. Una forma es el usado en el llamado máser de tres estados. Para éste, se usan sistemas atómicos que tienen tres niveles de energía, como se muestra en la figura 9-6, con las siguientes propiedades especiales. El sistema absorberá radiación (digamos, luz)

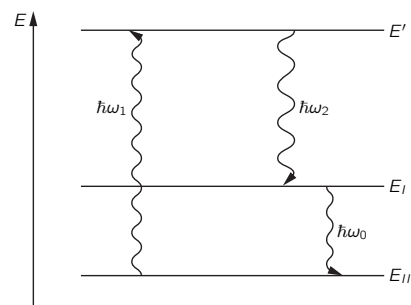


Figura 9-6. Niveles de energía en un máser de “tres estados”.

de frecuencia $\hbar\omega_1$ e irá del nivel más bajo de energía E_{II} , a un nivel de energía más alta E' , y luego emitirá rápidamente fotones de frecuencia $\hbar\omega_2$, pasando al estado $|I\rangle$ de energía E_I . El estado $|I\rangle$ tiene una vida larga, de modo que su población se puede aumentar, y así las condiciones son las apropiadas para un funcionamiento del máser entre los estados $|I\rangle$ y $|II\rangle$. Aunque tal dispositivo se llama máser de “tres estados”, el funcionamiento del máser se lleva a cabo realmente como en el sistema de dos estados que hemos estado describiendo.

Un láser (amplificación de la luz por emisión estimulada de la radiación en inglés) es simplemente un máser funcionando en frecuencias visibles. Por lo común, la “cavidad” de un láser consiste sólo en dos espejos planos, entre los cuales se generan ondas estacionarias.

9-5 Transiciones fuera de resonancia

Finalmente, queremos saber cómo varían los estados en el caso de que la frecuencia de la cavidad sea casi, pero no exactamente, igual a ω_0 . Podríamos resolver este problema de manera exacta, pero en lugar de tratar de hacerlo, tomaremos el caso importante en que el campo eléctrico es pequeño, y también el periodo de tiempo T es pequeño, por lo que $\mu\mathcal{E}_0T/\hbar$ es mucho menor que uno. Entonces, aun en el caso de resonancia perfecta que acabamos de ver, la probabilidad para hacer una transición es pequeña. Supongan que empezamos otra vez con $\gamma_I = 1$ y $\gamma_{II} = 0$. Durante el tiempo T esperaríamos que γ_I permaneciese cerca de uno, y que γ_{II} permaneciese muy pequeño comparado con la unidad. Así, el problema es muy fácil. Podemos calcular γ_{II} a partir de la segunda ecuación en (9.45), haciendo a γ_I igual a uno e integrando desde $t = 0$ hasta $t = T$. Obtenemos

$$\gamma_{II} = \frac{\mu\mathcal{E}_0}{\hbar} \left[\frac{1 - e^{i(\omega - \omega_0)T}}{\omega - \omega_0} \right]. \quad (9.51)$$

Este γ_{II} , usado junto con la ecuación (9.40), nos da la amplitud para haber efectuado una transición del estado $|I\rangle$ al $|II\rangle$ durante el intervalo de tiempo T . La probabilidad $P(I \rightarrow II)$ para efectuar la transición es $|\gamma_{II}|^2$, o sea

$$P(I \rightarrow II) = |\gamma_{II}|^2 = \left[\frac{\mu\mathcal{E}_0T}{\hbar} \right]^2 \frac{\text{sen}^2[(\omega - \omega_0)T/2]}{[(\omega - \omega_0)T/2]^2}. \quad (9.52)$$

Es interesante graficar esta probabilidad, para un intervalo fijo de tiempo, en función de la frecuencia de la cavidad, para ver cuán sensible es a las frecuencias cercanas a la frecuencia de resonancia ω_0 . Mostramos esa gráfica de $P(I \rightarrow II)$ en la figura 9-7. (La escala vertical se ha ajustado para que sea 1 en el máximo, dividiendo entre el valor de la probabilidad cuando $\omega = \omega_0$.) Vimos una curva como ésta en la teoría de difracción, por eso les debe resultar familiar. La curva cae abruptamente a cero para $(\omega - \omega_0) = 2\pi/T$ y nunca vuelve a tener tamaño significativo para grandes desviaciones de frecuencia. En efecto, la mayor parte del área bajo la curva se encuentra dentro del intervalo $\pm\pi/T$. Es posible demostrar† que el área debajo de la curva es sólo $2\pi/T$, y que es igual al área del rectángulo sombreado dibujado en la figura.

Examinemos las implicaciones de nuestros resultados para un máser real. Supongan que la molécula de amoníaco permanece en la cavidad un tiempo razonable, digamos un milisegundo. Entonces, para $f_0 = 24\,000$ megaciclos, podemos calcular que la probabilidad para una transición cae a cero para una desviación de frecuencia de $(f - f_0)/f_0 = 1/f_0T$, que es un quinto de 10^8 . Evidentemente, la frecuencia debe ser muy cercana a ω_0 para obtener una probabilidad de transición significativa. Tal efecto es la base de la gran precisión que se puede obtener con relojes “atómicos”, los cuales se basan en el principio del máser.

† Usando la fórmula $\int_{-\infty}^{\infty} (\text{sen}^2 x/x^2) dx = \pi$.

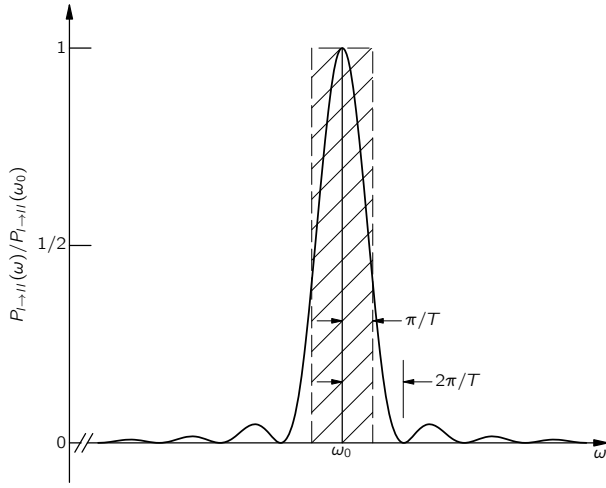


Figura 9-7. Probabilidad de transición para la molécula de amoníaco en función de la frecuencia.

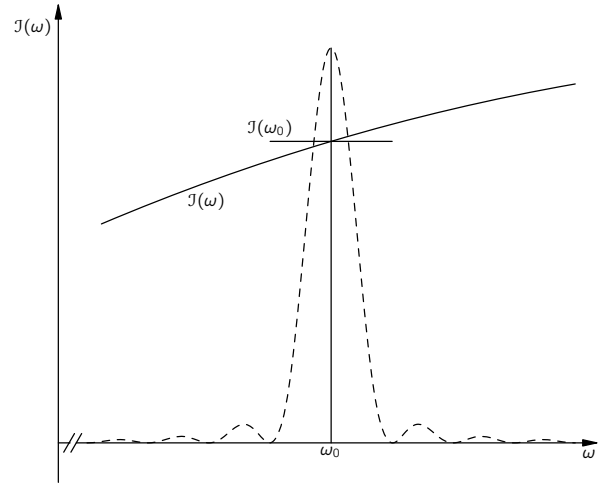


Figura 9-8. La intensidad espectral $J(\omega)$ se puede aproximar por su valor en ω_0 .

9-6 Absorción de luz

Nuestro procedimiento anterior se adapta a una situación más general que el máser de amoníaco. Hemos tratado el comportamiento de una molécula bajo la influencia de un campo eléctrico, ya el campo esté confinado a una cavidad o no. Así que podríamos simplemente iluminar la molécula con un haz de “luz” —con frecuencias de microondas— y preguntar por la probabilidad para emisión o absorción. Nuestras ecuaciones se adaptan igualmente bien a este caso, pero vamos a escribirlas en términos de la *intensidad* de radiación en lugar del campo eléctrico. Si definimos la intensidad $\mathcal{J}$ como el flujo de energía promedio por unidad de área por segundo, entonces, según el capítulo 27 del volumen II, podemos escribir

$$\mathcal{J} = \epsilon_0 c^2 |\mathcal{E} \times \mathbf{B}|_{\text{prom}} = \frac{1}{2} \epsilon_0 c^2 |\mathcal{E} \times \mathbf{B}|_{\text{máx}} = 2\epsilon_0 c \mathcal{E}_0^2.$$

(El valor máximo de $\mathcal{E}$ es $2\mathcal{E}_0$.) La probabilidad de transición ahora es:

$$P(I \rightarrow II) = 2\pi \left[\frac{\mu^2}{4\pi\epsilon_0\hbar^2 c} \right] \mathcal{J} T^2 \frac{\text{sen}^2[(\omega - \omega_0)T/2]}{[(\omega - \omega_0)T/2]^2}. \quad (9.53)$$

Por lo común, la luz que ilumina tal sistema no es exactamente monocromática. Por lo tanto, es interesante resolver otro problema —es decir, calcular la probabilidad de transición cuando la intensidad de la luz es $\mathcal{J}(\omega)$ por intervalo de frecuencia, cubriendo un intervalo grande que incluya a ω_0 —. Entonces, la probabilidad para ir de $|I\rangle$ a $|II\rangle$ se convertirá en una integral:

$$P(I \rightarrow II) = 2\pi \left[\frac{\mu^2}{4\pi\epsilon_0\hbar^2 c} \right] T^2 \int_0^\infty \mathcal{J}(\omega) \frac{\text{sen}^2[(\omega - \omega_0)T/2]}{[(\omega - \omega_0)T/2]^2} d\omega. \quad (9.54)$$

En general, $\mathcal{J}(\omega)$ variará con ω más lentamente que el término de resonancia. Las dos funciones pueden parecerse a las mostradas en la figura 9-8. En estos casos, podemos reemplazar $\mathcal{J}(\omega)$ por su valor $\mathcal{J}(\omega_0)$ en el centro de la curva de resonancia y sacarlo de la integral. Lo que queda es sólo la integral bajo la curva de la figura 9-7, que es, como hemos visto, igual a $2\pi/T$. Tenemos como resultado que

$$P(I \rightarrow II) = 4\pi^2 \left[\frac{\mu^2}{4\pi\epsilon_0\hbar^2 c} \right] \mathcal{J}(\omega_0) T. \quad (9.55)$$

Éste es un resultado importante, porque es la *teoría general de la absorción de la luz por cualquier sistema molecular o atómico*. Aunque comenzamos

considerando un caso en el que el estado $|I\rangle$ tenía una energía más alta que el estado $|II\rangle$, ninguno de nuestros argumentos depende de este hecho. La ecuación (9.55) es útil aunque el estado $|I\rangle$ tenga una energía *más baja* que el estado $|II\rangle$; entonces $P(I \rightarrow II)$ representa la probabilidad para una transición con *absorción* de energía de la onda electromagnética incidente. La absorción de luz por cualquier sistema atómico siempre implica la amplitud para una transición en un campo eléctrico oscilante, entre dos estados separados por una energía $E = \hbar\omega_0$. Para cualquier caso particular, se deduce siempre en la forma que lo hemos hecho aquí y da una expresión parecida a la ecuación (9.55). Por lo tanto, hacemos énfasis en las siguientes características de tal resultado. En primer lugar, la probabilidad es proporcional a T . En otras palabras, hay una probabilidad constante por unidad de tiempo para que ocurran las transiciones. En segundo lugar, dicha probabilidad es proporcional a la *intensidad* de la luz incidente sobre el sistema. Finalmente, la probabilidad de transición es proporcional a μ^2 , donde, como recordarán, $\mu\mathcal{E}$ definía el cambio de energía debido al campo eléctrico $\mathcal{E}$. Por lo anterior, $\mu\mathcal{E}$ también aparecía en las ecuaciones (9.38) y (9.39) como el término de acoplamiento responsable de la transición entre los estados $|I\rangle$ y $|II\rangle$, que de otra manera son estacionarios. En otras palabras, para el pequeño $\mathcal{E}$ que hemos estado considerando, $\mu\mathcal{E}$ es el llamado “término de perturbación” en el elemento de la matriz hamiltoniana, este término conecta los estados $|I\rangle$ y $|II\rangle$. En el caso general, tenemos que $\mu\mathcal{E}$ es remplazado por el elemento matricial $\langle II | H | I \rangle$ (vean la sección 5-6).

En el volumen I (sección 42-5) hablamos de las relaciones entre la absorción de luz, la emisión inducida y la emisión espontánea en términos de los coeficientes A y B de Einstein. Aquí, por fin tenemos el procedimiento cuántico para calcular estos coeficientes. Lo que hemos llamado $P(I \rightarrow II)$ para nuestra molécula de amoníaco con dos estados, corresponde precisamente al coeficiente de absorción B_{nm} de la teoría de radiación de Einstein. Para la complicada molécula de amoníaco —que es muy difícil de calcular— hemos tomado el elemento matricial $\langle II | H | I \rangle$ igual a $\mu\mathcal{E}$, diciendo que μ se tiene que obtener experimentalmente. Para sistemas atómicos más simples, la μ_{mn} que pertenece a cualquier transición particular se puede calcular de la *definición*

$$\mu_{mn}\mathcal{E} = \langle m | H | n \rangle = H_{mn}, \quad (9.56)$$

donde H_{mn} es el elemento matricial del hamiltoniano que incluye los efectos de un campo eléctrico débil. La μ_{mn} calculada de esta forma, se le llama el *elemento matricial de dipolo eléctrico*. La teoría cuántica de la absorción y la emisión de luz es, entonces, reducida al cálculo de estos elementos matriciales para sistemas atómicos determinados.

Nuestro estudio de un sistema sencillo de dos estados nos han conducido a una comprensión del problema general de la absorción y la emisión de luz.

Otros sistemas de dos estados

10-1 El ion del hidrógeno molecular

En el capítulo anterior estudiamos algunos aspectos de la molécula de amoníaco, con la aproximación de que ésta se puede considerar como un sistema de dos estados. Por supuesto, no es realmente un sistema de dos estados —hay muchos estados de rotación, de vibración, de traslación, etc.— pero cada uno de estos estados de movimiento se debe analizar en términos de dos estados internos, debidos al cambio de posición del átomo de nitrógeno. Ahora vamos a tratar otros ejemplos de sistemas que, con una u otra aproximación, se pueden considerar como sistemas de dos estados. Muchas cosas serán aproximadas, porque siempre hay muchos otros estados, que en un análisis más preciso habría que tener en cuenta. No obstante, en cada ejemplo nos será posible comprender muchas cosas considerando sólo dos estados.

Como trataremos solamente con sistemas de dos estados, el hamiltoniano que necesitamos será como el que usamos en el capítulo anterior. Cuando el hamiltoniano es independiente del tiempo, sabemos que hay dos estados estacionarios con energía definida —en general diferente—. Sin embargo, comúnmente comenzamos nuestro análisis con un conjunto de estados base que *no* son estos estados estacionarios, sino estados que pueden, tal vez, tener otro significado físico simple. Entonces, los estados estacionarios del sistema estarán representados por una combinación lineal de estos estados base.

Por conveniencia, resumiremos las ecuaciones más importantes del capítulo 9. Si los estados base originales son $|1\rangle$ y $|2\rangle$. Entonces cualquier estado $|\psi\rangle$ está representado por la combinación lineal

$$|\psi\rangle = |1\rangle\langle 1|\psi\rangle + |2\rangle\langle 2|\psi\rangle = |1\rangle C_1 + |2\rangle C_2. \quad (10.1)$$

Las amplitudes C_i (con lo que nos referimos a C_1 o C_2) satisfacen las dos ecuaciones diferenciales lineales:

$$i\hbar \frac{dC_i}{dt} = \sum_j H_{ij} C_j, \quad (10.2)$$

donde ambos i y j toman los valores 1 y 2.

Cuando los términos del hamiltoniano H_{ij} no dependen de t , los dos estados de energía definida (estados estacionarios), que llamamos

$$|\psi_I\rangle = |I\rangle e^{-(i/\hbar)E_I t} \quad \text{y} \quad |\psi_{II}\rangle = |II\rangle e^{-(i/\hbar)E_{II} t},$$

tienen las energías

$$E_I = \frac{H_{11} + H_{22}}{2} + \sqrt{\left(\frac{H_{11} - H_{22}}{2}\right)^2 + H_{12}H_{21}},$$

$$E_{II} = \frac{H_{11} + H_{22}}{2} - \sqrt{\left(\frac{H_{11} - H_{22}}{2}\right)^2 + H_{12}H_{21}}. \quad (10.3)$$

Las dos C de cada uno de estos estados tienen la misma dependencia temporal. Los vectores de estado $|I\rangle$ y $|II\rangle$, que aparecen en los dos estados estacionarios, están relacionados con los estados base originales $|1\rangle$ y $|2\rangle$ mediante

$$|I\rangle = |1\rangle a_1 + |2\rangle a_2,$$

$$|II\rangle = |1\rangle a'_1 + |2\rangle a'_2. \quad (10.4)$$

10-1 El ion del hidrógeno molecular

10-2 Fuerzas nucleares

10-3 La molécula de hidrógeno

10-4 La molécula de benceno

10-5 Colorantes

10-6 El hamiltoniano de una partícula con espín un medio en un campo magnético

10-7 El electrón con espín en un campo magnético

Las a son constantes complejas, las cuales satisfacen

$$|a_1|^2 + |a_2|^2 = 1,$$

$$\frac{a_1}{a_2} = \frac{H_{12}}{E_I - H_{11}}, \quad (10.5)$$

$$|a'_1|^2 + |a'_2|^2 = 1,$$

$$\frac{a'_1}{a'_2} = \frac{H_{12}}{E_{II} - H_{11}}. \quad (10.6)$$

Si H_{11} y H_{22} son iguales —digamos que son iguales a E_0 — y $H_{12} = H_{21} = -A$, entonces $E_I = E_0 + A$, $E_{II} = E_0 - A$ y los estados $|I\rangle$ y $|II\rangle$ son particularmente simples:

$$|I\rangle = \frac{1}{\sqrt{2}} [|1\rangle - |2\rangle], \quad |II\rangle = \frac{1}{\sqrt{2}} [|1\rangle + |2\rangle]. \quad (10.7)$$

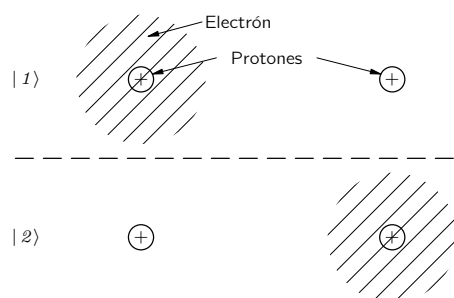


Figura 10-1. Un conjunto de estados base para dos protones y un electrón.

Ahora utilizaremos estos resultados para estudiar algunos ejemplos interesantes tomados de la química y de la física. El primer ejemplo es el ion del hidrógeno molecular. Una molécula de hidrógeno ionizada positivamente consiste en dos protones y un electrón girando a su alrededor. Si los dos protones están muy alejados uno de otro, ¿qué estados se esperarían en este sistema? La respuesta es bastante clara: el electrón se quedará cerca de un protón, formando un átomo de hidrógeno en su estado más bajo, y el otro protón quedará solo como ion positivo. Así, si los dos protones están alejados, podemos imaginar un estado físico en el que el electrón está “adherido” a uno de los protones. Hay, evidentemente, otro estado simétrico al anterior, en el cual el electrón está cerca del otro protón y el primer protón es el ion. Tomaremos a estos dos como estado base y los llamaremos $|1\rangle$ y $|2\rangle$. Están representados en la figura 10-1. Por supuesto, en realidad hay muchos estados de un electrón cerca de un protón, porque la combinación puede existir como cualquiera de los estados excitados del átomo de hidrógeno. Ahora no estamos interesados en esa variedad de estados; sólo consideraremos el caso del átomo de hidrógeno en su estado más bajo —su estado base— y, por el momento, dejaremos de lado el espín del electrón. Suponemos simplemente que, en todos nuestros estados, el electrón tiene su espín “hacia arriba” en la dirección del eje z .†

Ahora bien, para quitar un electrón de un átomo de hidrógeno se necesitan 13.6 electronvoltios de energía. Mientras los dos protones del ion del hidrógeno molecular estén alejados, aún se necesita más o menos esta energía —que es mucha energía para lo que estamos considerando por el momento— para llevar al electrón cerca del punto medio entre los dos protones. Por consiguiente, es imposible, de forma clásica, que el electrón salte de un protón al otro. Sin embargo, en mecánica cuántica es posible —aunque no muy probable—. Hay una pequeña amplitud para que el electrón pase de un protón al otro. Como primera aproximación, cada uno de los estados base $|1\rangle$ y $|2\rangle$ tendrá la energía E_0 , que es justamente la energía de un átomo de hidrógeno más un protón. Podemos considerar que los elementos de la matriz hamiltoniana, H_{11} y H_{22} , son aproximadamente iguales a E_0 . De nuevo, tomaremos los otros elementos matriciales H_{12} y H_{21} , que son las amplitudes para que el electrón cambie de lugar, iguales a $-A$.

Como pueden observar, es el mismo procedimiento que estuvimos practicando en los dos capítulos anteriores. Si ignoramos el hecho de que el electrón puede cambiar de lugar, tenemos dos estados de energía exactamente igual. Sin embargo, esta energía se separará en dos niveles de energía debido a la posibilidad de que el electrón cambie de lugar —cuanto mayor es la probabilidad para la transición,

† Esto es satisfactorio mientras no haya campos magnéticos importantes. Más adelante en este capítulo estudiaremos los efectos de un campo magnético sobre el electrón, también estudiaremos los muy pequeños efectos del espín en el átomo de hidrógeno en el capítulo 12.

mayor es la separación—. Así, los dos niveles de energía del sistema son $E_0 + A$ y $E_0 - A$; y los estados que tienen estas energías definidas están dados por las ecuaciones (10.7).

De acuerdo con nuestra solución, vemos que, si se pone a un protón y a un átomo de hidrógeno muy juntos, el electrón no permanecerá en uno de los protones, sino que saltará de un protón a otro. Si el electrón comienza en uno de los protones, oscilará entre los estados $|I\rangle$ y $|II\rangle$ —lo cual da una solución que varía con el tiempo—. Para obtener la solución de energía más baja (que no varía en el tiempo), es necesario que el sistema comience con amplitudes iguales para que el electrón esté alrededor de cada protón. Recuerden que no hay dos electrones —no estamos diciendo que hay un electrón alrededor de cada protón—. Hay solamente *un* electrón, y *éste* tiene la misma amplitud $1/\sqrt{2}$ en magnitud— para estar en cualquier posición.

Ahora bien, la amplitud A para que un electrón que está cerca de un protón vaya hasta el otro, depende de la distancia entre los protones. Cuanto más cerca estén los protones, mayor será la amplitud. Recuerden que en el capítulo 7 hablamos de la amplitud para que un electrón “atravesara una barrera”, lo cual era imposible de hacer de manera clásica. Aquí tenemos la misma situación. La amplitud para que un electrón atravesara disminuye en forma aproximadamente exponencial con la distancia —para distancias grandes—. Como la probabilidad para la transición, y por lo tanto A , aumenta cuando los protones se acercan, también aumentará la separación entre los niveles de energía. Si el sistema está en el estado $|I\rangle$, la energía $E_0 + A$ aumenta al disminuir la distancia, por lo que estos efectos cuánticos equivalen a una fuerza *repulsiva* que tiende a mantener alejados a los protones. Por otra parte, si el sistema está en el estado $|II\rangle$, la energía total *disminuye* si los protones se acercan; hay una fuerza *atractiva* que intenta acercar a los protones. La variación de las dos energías con la distancia entre los protones debe ser, más o menos, como muestra la figura 10-2. Tenemos, entonces, una explicación cuántica de la fuerza de enlace que mantiene unido al ion H_2^+ .

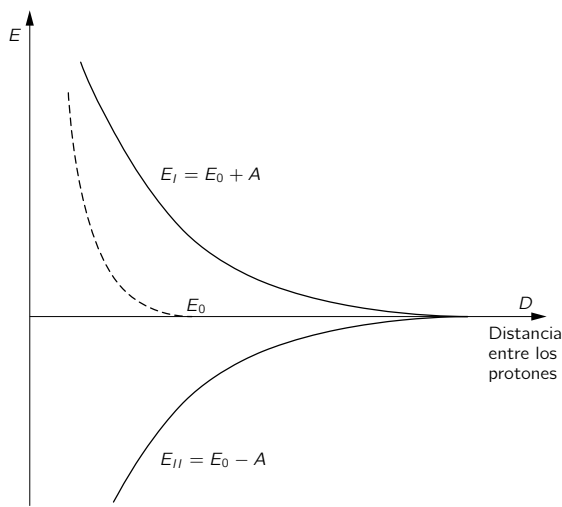


Figura 10-2. Las energías de los dos estados estacionarios del ion H_2^+ en función de la distancia entre los dos protones.

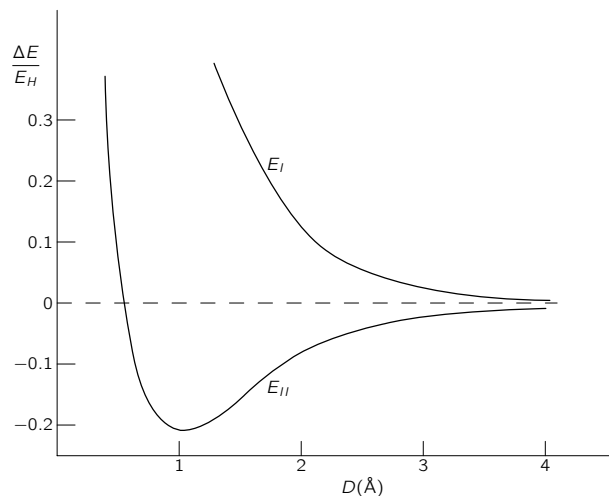


Figura 10-3. Los niveles de energía del ion H_2^+ en función de la distancia interprotónica D . ($E_H = 13.6\text{eV}$.)

Pero nos hemos olvidado de algo. Además de la fuerza que acabamos de describir, también hay una fuerza electrostática repulsiva entre los dos protones. Cuando los dos protones están muy alejados —como en la figura 10-1— el protón “descubierto” sólo ve un átomo neutro, por lo que la fuerza electrostática es despreciable. A distancias muy pequeñas, sin embargo, el protón “descubierto”

comienza a “meterse” en la distribución electrónica —es decir, que en promedio, está más cerca del protón que del electrón—. Así, empieza a producirse una cierta energía electrostática adicional que, naturalmente, es positiva. Esta energía —que también varía con la distancia— debe incluirse en E_0 . Por lo tanto, para E_0 , debemos tomar algo así como la línea discontinua de la figura 10-2, la cual crece rápidamente para distancias menores que el radio del átomo de hidrógeno. Debemos sumar y restar la energía correspondiente a los cambios de lugar A de esta E_0 . Si hacemos esto, las energías E_I y E_{II} variarán con la distancia interprotónica D como se muestra en la figura 10-3. [En esta figura, hemos graficado los resultados de un cálculo más detallado. La distancia interprotónica está dada en unidades de $1\text{\AA}$ (10^{-8} cm), y la energía en exceso sobre la de un protón más un átomo de hidrógeno está dada en unidades de la energía de enlace del átomo de hidrógeno —la llamada energía de “Rydberg”, 13.6 eV —.] Vemos que el estado $|II\rangle$ tiene un punto de energía mínima. Éste corresponde a la configuración de equilibrio —el estado de energía más baja— del ion H_2^+ . En este punto, la energía es más baja que la de un protón y un átomo de hidrógeno separados, por lo que el sistema está enlazado. Un solo electrón funciona para mantener unidos a los dos protones. Un químico lo llamaría un “enlace de un electrón”.

Este tipo de unión química también se denomina frecuentemente “resonancia cuántica” (por analogía con los dos péndulos acoplados que describimos antes). Pero en realidad parece más misterioso de lo que es; sólo es una “resonancia” si comienzan con una mala elección de los estados base —¡como nosotros hicimos!—. Si hubieran elegido el estado $|II\rangle$, tendrían al estado de energía más baja —eso es todo—.

Hay otra manera de ver por qué ese estado debe tener una energía más baja que la de un protón y un átomo de hidrógeno. Imaginemos a un electrón cerca de dos protones a una distancia fija, pero no muy grande. Recordarán que con un solo protón el electrón está “disperso” debido al principio de incertidumbre. Busca un equilibrio entre tener una energía *potencial* de Coulomb baja y no quedarse confinado en un espacio demasiado pequeño, lo cual provocaría una energía *cinética* alta (debido a la relación de incertidumbre $\Delta p \Delta x \approx \hbar$). Ahora bien, cuando hay dos protones, hay más espacio donde el electrón puede tener una energía potencial baja. Se puede dispersar —disminuyendo su energía cinética— sin aumentar su energía potencial. El resultado efectivo es una energía más baja que la de un protón y un átomo de hidrógeno. Entonces, ¿por qué el otro estado $|I\rangle$ tiene energía más alta? Observen que este estado es la *diferencia* de los estados $|1\rangle$ y $|2\rangle$. Debido a la simetría de $|1\rangle$ y $|2\rangle$, la diferencia debe dar una amplitud cero para encontrar al electrón a la mitad entre los dos protones. Esto significa que el electrón está un poco más confinado, lo cual da lugar a una energía mayor.

Debemos señalar que nuestro tratamiento aproximado del ion H_2^+ como un sistema de dos estados, falla estrepitosamente cuando los protones se acercan tanto como lo están en el mínimo de la curva de la figura 10-3, por lo cual, no obtendremos un buen valor para la energía de enlace real. Para distancias pequeñas, las energías de los dos “estados” que imaginamos en la figura 10-1 no son, en realidad, iguales a E_0 ; se necesita un tratamiento cuántico más refinado.

Ahora, nos preguntamos qué ocurriría si, en vez de dos protones, tuviéramos dos objetos diferentes —como, por ejemplo, un protón y un ion positivo de litio (ambas partículas tienen una sola carga positiva)—. En este caso, los dos términos H_{11} y H_{22} del hamiltoniano ya no serán iguales; en realidad, serán muy diferentes. Si ocurriera que la diferencia ($H_{11} - H_{22}$) fuese, en valor absoluto, mucho mayor que $A = -H_{12}$, la fuerza atractiva se haría muy débil, como se puede ver de la siguiente manera.

Si hacemos $H_{12}H_{21} = A^2$ en las ecuaciones (10.3) obtenemos

$$E = \frac{H_{11} + H_{22}}{2} \pm \frac{H_{11} - H_{22}}{2} \sqrt{1 + \frac{4A^2}{(H_{11} - H_{22})^2}}.$$

Cuando $H_{11} - H_{22}$ es mucho mayor que A^2 , la raíz cuadrada es cercana a

$$1 + \frac{2A^2}{(H_{11} - H_{22})^2}.$$

Las dos energías son entonces

$$\begin{aligned} E_I &= H_{11} + \frac{A^2}{(H_{11} - H_{22})}, \\ E_{II} &= H_{22} - \frac{A^2}{(H_{11} - H_{22})}. \end{aligned} \quad (10.8)$$

Ahora son casi iguales las energías H_{11} y H_{22} de los átomos aislados, apartadas sólo ligeramente por la amplitud A para cambiar de lugar.

La diferencia de energía $E_I - E_{II}$ es

$$(H_{11} - H_{22}) + \frac{2A^2}{H_{11} - H_{22}}.$$

La separación *adicional* causada por el cambio de lugar del electrón ya no es igual a $2A$; está disminuida en un factor de $A/(H_{11} - H_{22})$, que estamos suponiendo mucho menor que uno. Además, la dependencia de $E_I - E_{II}$ con la separación de los dos núcleos es mucho menor que para el ion H_2^+ —también se reducen en un factor de $A/(H_{11} - H_{22})$ —. Ahora podemos ver por qué el enlace en moléculas diatómicas asimétricas es, generalmente, muy débil.

En nuestra teoría sobre el ion H_2^+ , hemos descubierto una explicación para el mecanismo por el cual un electrón compartido por dos protones causa una fuerza de atracción entre ambos protones, la cual puede existir aun cuando los dos protones estén alejados por grandes distancias. La fuerza atractiva proviene de la reducción de energía del sistema, debida a la posibilidad de que el electrón salte de un protón al otro. En esos saltos, el sistema cambia de la configuración (átomo de hidrógeno, protón) a la configuración (protón, átomo de hidrógeno), o al revés. Podemos escribir simbólicamente el proceso en la forma

$$(H, p) \rightleftharpoons (p, H).$$

El cambio de energía debido a este proceso es proporcional a la amplitud A para que un electrón de energía $-W_H$ (su energía de unión en el átomo de hidrógeno) pueda pasar de un protón al otro.

Para grandes distancias R entre ambos protones, la energía potencial electrostática del electrón es casi cero en la mayor parte del espacio que debe recorrer al efectuar el salto. En este espacio, por lo tanto, el electrón se mueve casi como una partícula libre en el espacio vacío —¡pero con energía *negativa*!—. Hemos visto en el capítulo 3 [ecuación (3.7)] que la amplitud para que una partícula de energía definida vaya de un lugar a otro, que está a una distancia r , es proporcional a

$$\frac{e^{(i/\hbar)pr}}{r},$$

donde p es el momentum correspondiente a esa energía definida. En el caso actual (usando la fórmula no relativista), p está dado por

$$\frac{p^2}{2m} = -W_H. \quad (10.9)$$

Esto significa que p es un número imaginario,

$$p = i\sqrt{2mW_H}$$

(el otro signo de la raíz no tiene sentido aquí).

Entonces, debemos esperar que la amplitud A para el ion H_2^+ varíe en la forma

$$A \propto \frac{e^{-(\sqrt{2mW_H}/\hbar)R}}{R} \quad (10.10)$$

para distancias grandes R entre ambos protones. El cambio de energía debido al enlace del electrón es proporcional a A , por lo que hay una fuerza tratando de acercar los dos protones, la cual es proporcional —para R grandes— a la derivada de (10.10) respecto a R .

Finalmente, para completar, debemos señalar que, en el sistema de dos protones y un electrón, hay aún otro efecto que da una dependencia para la energía con R . Lo hemos despreciado hasta el momento porque normalmente no es muy importante —la excepción se presenta precisamente para distancias muy grandes, en las cuales la energía del término de intercambio A ha disminuido exponencialmente hasta valores muy pequeños—. El nuevo efecto que estamos considerando es la atracción electrostática del protón hacia el átomo de hidrógeno, la cual ocurre de la misma manera en que un objeto cargado atrae un objeto neutro. El protón descubierto establece un campo eléctrico $\mathcal{E}$ (que varía como $1/R^2$) hacia el átomo de hidrógeno neutro. El átomo se polariza, adquiriendo un momento dipolar inducido μ proporcional a $\mathcal{E}$. La energía del dipolo es $\mu\mathcal{E}$, que es proporcional a $\mathcal{E}^2$ —es decir, a $1/R^4$ —. Así, hay un término en la energía del sistema que decrece con la potencia cuarta de la distancia. (Es una corrección a E_0 .) Esta energía cae más lentamente con la distancia que el cambio A dado por (10.10); a cierta distancia R suficientemente grande, es el único término importante que sigue dando una variación de la energía con R —y, por lo tanto, es la única fuerza que queda—. Observen que el término electrostático tiene el mismo signo para los dos estados base (la fuerza es atractiva, por lo que la energía es negativa) y así también para los dos estados estacionarios, mientras que el término para el intercambio del electrón A tiene signos opuestos para los dos estados estacionarios.

10-2 Fuerzas nucleares

Hemos visto que el sistema de un átomo de hidrógeno y un protón tiene una energía de interacción debida al intercambio de un único electrón, la cual varía para grandes distancias R como

$$\frac{e^{-\alpha R}}{R}, \quad (10.11)$$

con $\alpha = \sqrt{2mW_H}/\hbar$. (Normalmente se dice que hay un intercambio de un electrón “virtual” cuando —como aquí— el electrón tiene que atravesar una región donde tendría una energía negativa. Más específicamente, un “intercambio virtual” significa que el fenómeno implica una interferencia cuántica entre un estado intercambiado y uno no intercambiado.)

Ahora, podríamos hacer la siguiente pregunta: ¿es posible que fuerzas entre otros tipos de partículas tengan un origen análogo? ¿Qué pasa, por ejemplo, con la fuerza nuclear entre un neutrón y un protón, o entre dos protones? Al intentar explicar la naturaleza de las fuerzas nucleares, Yukawa propuso que la fuerza entre dos nucleones se debe a un efecto similar de intercambio —sólo que, en este caso, se debe al intercambio virtual, no de un electrón, sino de una nueva partícula que denominó “mesón”. Actualmente, identificamos el mesón de Yukawa con el mesón π (o “pion”) producido en las colisiones de altas energías entre protones u otras partículas.

Veamos, como ejemplo, qué tipo de fuerza esperaríamos del intercambio de un pion positivo (π^+) de masa m_π entre un protón y un neutrón. Como un átomo de hidrógeno H^0 puede transformarse en un protón p^+ liberando un electrón e^- ,

$$H^0 \rightarrow p^+ + e^-, \quad (10.12)$$

un protón p^+ puede transformarse en un neutrón n^0 liberando un mesón π^+ :

$$p^+ \rightarrow n^0 + \pi^+. \quad (10.13)$$

Así, si tenemos un protón en a y un neutrón en b separados por una distancia R , el protón se puede convertir en un neutrón emitiendo un π^+ , que luego es

absorbido por el neutrón en b , convirtiéndolo en un protón. Hay una energía de interacción en el sistema de dos nucleones (más el pion) que depende de la amplitud A para que se intercambie un pion —como vimos en el intercambio de un electrón en el ion H_2^+ —.

En el proceso (10.12), la energía del átomo de H^0 es menor que la energía del protón por W_H (calculando en forma no relativista, y omitiendo la energía en reposo mc^2 del electrón), de modo que el electrón tiene energía *cinética* negativa —o sea un momentum imaginario— como en la ecuación (10.9). En el proceso nuclear (10.13), el protón y el neutrón tienen masas casi iguales, por lo que el π^+ tendrá energía *total* cero. La relación entre la energía total E y el momentum p para un pion de masa m_π es

$$E^2 = p^2 c^2 + m_\pi^2 c^4.$$

Como E es cero (o por lo menos despreciable en comparación con m_π), también aquí el momentum es imaginario:

$$p = im_\pi c.$$

Utilizando los mismos argumentos que dimos para la amplitud para que un electrón enlazado atraviese la barrera en el espacio entre dos protones, obtenemos, para el caso nuclear, una amplitud de intercambio A que —para R grande— tiene que variar como

$$\frac{e^{-(m_\pi c/\hbar)R}}{R}. \quad (10.14)$$

La energía de interacción es proporcional a A , por lo que varía del mismo modo. Obtenemos una variación de energía de la forma del denominado *potencial de Yukawa* entre dos nucleones. Además, obtuvimos esta misma fórmula antes, a partir de la ecuación diferencial para el movimiento de un pion en el espacio libre [véase volumen II, capítulo 28, ecuación (28.18)].

Podemos, con el mismo argumento, discutir la interacción entre dos protones (o entre dos neutrones) que resulta del intercambio de un pion neutro (π^0). El proceso básico es ahora

$$p^+ \rightarrow p^+ + \pi^0. \quad (10.15)$$

Un protón puede emitir un π^0 virtual, pero sigue siendo un protón. Si tenemos dos protones, el protón núm. 1 puede emitir un π^0 virtual, el cual es absorbido por el protón núm. 2. Al final del proceso, aún tenemos dos protones. Esto es un tanto distinto al ion H_2^+ . Ahí el H^0 pasaba a un estado diferente —al protón— después de emitir el electrón. Ahora estamos suponiendo que un protón puede emitir un π^0 sin cambiar su carácter. Dichos procesos se observan realmente en colisiones de alta energía. El proceso es análogo a la manera en que un electrón emite un fotón y sigue siendo un electrón:

$$e \rightarrow e + \text{fotón}. \quad (10.16)$$

No “vemos” a los fotones dentro de los electrones antes de que sean emitidos o después que son absorbidos, y su emisión no altera la “naturaleza” del electrón.

Volviendo a los dos protones, hay una energía de interacción que proviene de la amplitud A para que un protón emita un pion neutro, que viaja (con momentum imaginario) hasta el otro protón y es absorbido. Esta amplitud es proporcional a (10.14), donde m_π es la masa del pion neutro. Los mismos argumentos dan una energía de interacción igual para dos neutrones. Como las fuerzas nucleares (ignorando los efectos eléctricos) entre un neutrón y un protón, entre un protón y un protón, y entre un neutrón y un neutrón son iguales, concluimos que las masas de los piones cargados y neutros deben ser iguales. Experimentalmente, las masas son casi iguales, y la pequeña diferencia está cerca de lo que se espera por las correcciones de autoenergía eléctrica (véase capítulo 28, volumen II).

Hay otras clases de partículas —como los mesones K — que se pueden intercambiar entre dos nucleones. También es posible que se intercambien dos piones

al mismo tiempo. Pero estos otros “objetos” intercambiados tienen una masa en reposo m_x mayor que la masa m_π del pion, y dan lugar a términos en la amplitud de intercambio que varían como

$$\frac{e^{-(m_x c/\hbar)R}}{R}.$$

Estos términos desaparecen más rápidamente al aumentar R que el término correspondiente a un mesón. Nadie sabe, actualmente, cómo calcular estos términos correspondientes a masas más altas, pero para valores de R suficientemente grandes sólo sobrevive el término correspondiente a un pion. Y de hecho, los experimentos que involucran interacciones nucleares sólo a grandes distancias, muestran que la energía de interacción es como la predicha por la teoría de intercambio de un pion.

En la teoría clásica de la electricidad y el magnetismo, la interacción electrostática de Coulomb y la radiación de luz por una carga acelerada están estrechamente relacionadas —ambas surgen de las ecuaciones de Maxwell—. Hemos visto en la teoría cuántica, que se puede representar a la luz como excitaciones cuánticas de las oscilaciones armónicas de los campos electromagnéticos clásicos dentro de una caja. Alternativamente, la teoría cuántica se puede establecer describiendo a la luz en términos de partículas —fotones— que siguen la estadística de Bose. En la sección 4-5 remarcamos que las dos alternativas siempre dan predicciones idénticas. ¿Es posible que el segundo punto de vista se pueda desarrollar para incluir *todos* los efectos electromagnéticos? En particular, si queremos describir al campo electromagnético sólo en términos de bosones —es decir, en términos de fotones— ¿a qué se debe la fuerza de Coulomb?

Desde el punto de vista de “partículas”, la interacción de Coulomb entre dos electrones *proviene del intercambio de un fotón virtual*. Un electrón emite un fotón —como en la reacción (10.16)— que va hacia el segundo electrón, donde es absorbido según la dirección inversa de la misma reacción. La energía de interacción está dada de nuevo por una fórmula como la (10.14), sólo que ahora hay que remplazar m_π por la masa en reposo del fotón —que es cero—. Así que el intercambio virtual de un fotón entre dos electrones da una energía de interacción que varía simplemente como la inversa de R , la distancia entre ambos electrones —exactamente la energía potencial de Coulomb!—. En la teoría de “partículas” del electromagnetismo, el proceso de intercambio de un fotón virtual da lugar a todos los fenómenos de la electrostática.

10-3 La molécula de hidrógeno

El siguiente sistema de dos estados que examinaremos es la molécula neutra de hidrógeno H_2 . Naturalmente, es más complicada de entender porque tiene dos electrones. De nuevo, comenzamos considerando qué ocurre cuando los dos protones están muy separados. Sólo que ahora tenemos que agregar dos electrones. Para seguirlos llamaremos “electrón a ” a uno de ellos y “electrón b ” al otro. De nuevo podemos imaginar dos estados posibles. Una posibilidad es que el “electrón a ” esté alrededor del primer protón y el “electrón b ” esté alrededor del segundo, como se muestra en la parte superior de la figura 10-4. Tenemos simplemente dos átomos de hidrógeno. Llamaremos a este estado $|1\rangle$. También hay otra posibilidad: que el “electrón b ” esté alrededor del primer protón y el “electrón a ” esté alrededor del segundo. Llamemos a este estado $|2\rangle$. Por la simetría de la situación, esas dos posibilidades deben ser equivalentes en términos de energía, pero como veremos, la energía del sistema *no* es simplemente la energía de dos átomos de hidrógeno. Tenemos que mencionar que hay muchas otras posibilidades. Por ejemplo, el “electrón a ” podría estar cerca del primer protón y el “electrón b ” podría estar en otro estado alrededor *del mismo* protón. Descartaremos este caso, ya que tendrá ciertamente una energía mayor (por la gran repulsión de Coulomb entre los dos electrones). Para mayor precisión, tendríamos que incluir dichos estados, pero podemos obtener lo esencial de la unión molecular considerando solamente los dos estados de la figura 10-4. En esta aproximación podemos

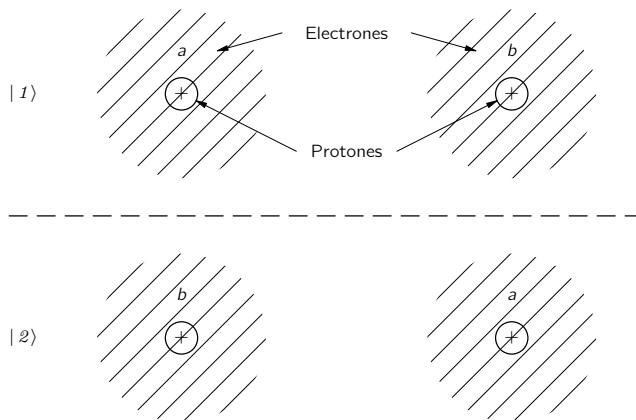


Figura 10-4. Un conjunto de estados base para la molécula de H_2 .

describir cualquier estado dando la amplitud $\langle 1 | \phi \rangle$ para estar en el estado $| 1 \rangle$ y la amplitud $\langle 2 | \phi \rangle$ para estar en el estado $| 2 \rangle$. En otras palabras, el vector de estado $|\phi\rangle$ se puede escribir como una combinación lineal

$$|\phi\rangle = \sum_i |i\rangle \langle i | \phi \rangle.$$

Para continuar, suponemos —como de costumbre— que hay cierta amplitud A para que los electrones se puedan mover por el espacio intermedio e intercambiar lugares. Esta posibilidad de intercambio significa que la energía del sistema se separa, como hemos visto en otros sistemas de dos estados. En lo que respecta al ion de hidrógeno molecular, la separación es muy pequeña cuando la distancia entre los protones es grande. A medida que los protones se aproximan, la amplitud para que los electrones se intercambien se incrementa, por lo que aumenta la separación. El decrecimiento del estado de energía menor significa que hay una fuerza atractiva que tiende a acercar los átomos. Nuevamente, los niveles de energía aumentan cuando los protones se acercan aún más, debido a la repulsión de Coulomb. El resultado final es que los dos estados estacionarios tienen energías que varían con la separación como se muestra en la figura 10-5. A una separación de unos $0.74 \text{ \AA}$, el nivel de energía menor alcanza un mínimo; ésta es la distancia protón-protón de la verdadera molécula de hidrógeno.

Probablemente ustedes hayan estado pensando en una objeción. ¿Y qué pasa con el hecho de que los dos electrones son partículas idénticas? Los hemos estado llamando “electrón a ” y “electrón b ”, pero realmente no hay manera de distinguir cuál es cuál. Y hemos dicho en el capítulo 4 que, para electrones —que son fermiones— si hay dos maneras para que algo pueda ocurrir al intercambiar dos electrones, las dos amplitudes interferirán con signo *negativo*. Esto significa que si intercambiamos cuál electrón es cuál, el signo de la amplitud se debe invertir. Acabamos de concluir, no obstante, que el estado enlazado de la molécula de hidrógeno sería (a $t = 0$)

$$|II\rangle = \frac{1}{\sqrt{2}} (|1\rangle + |2\rangle).$$

Sin embargo, según las reglas que dimos en el capítulo 4, este estado no está permitido. Si intercambiamos cuál electrón es cuál, obtenemos el estado

$$\frac{1}{\sqrt{2}} (|2\rangle + |1\rangle).$$

y tenemos el mismo signo en vez del opuesto.

Estos argumentos son correctos *si ambos electrones tienen el mismo espín*. Es cierto que cuando ambos electrones tienen el espín hacia arriba (o ambos hacia abajo), el único estado permitido es

$$|I\rangle = \frac{1}{\sqrt{2}} (|1\rangle - |2\rangle).$$

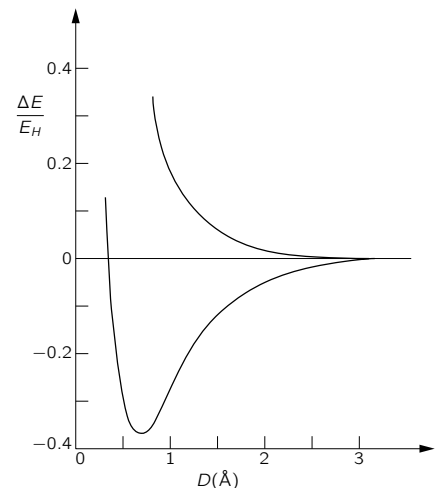


Figura 10-5. Los niveles de energía de la molécula H_2 para diferentes distancias interprotónicas D . ($E_H = 13.6 \text{ eV}$.)

Para este estado, el intercambio de los dos electrones da

$$\frac{1}{\sqrt{2}}(|2\rangle - |1\rangle).$$

que es $-|I\rangle$, como debe ser. Por lo que, si acercamos dos átomos de hidrógeno cuyos electrones tienen el espín en la misma dirección, pueden ir al estado $|I\rangle$ y no al estado $|II\rangle$. Pero observen que el estado $|I\rangle$ es el estado de energía *mayor*. Su curva de energía en función de la separación no tiene mínimo. Los dos hidrógenos se repelerán siempre y no formarán una molécula. Por lo que concluimos que la molécula de hidrógeno no puede existir con espines electrónicos paralelos. Y esto es correcto.

Por otra parte, nuestro estado $|II\rangle$ es perfectamente simétrico para los dos electrones. De hecho, si intercambiamos cuál es el electrón que llamamos *a* y cuál es el que llamamos *b* volvemos a obtener exactamente el mismo estado. Vimos en la sección 4-7 que si dos fermiones están en el mismo estado, *deben* tener espines opuestos. Por lo tanto, la molécula de hidrógeno enlazada debe tener un electrón con espín hacia arriba y el otro con espín hacia abajo.

El asunto completo de la molécula de hidrógeno es, en realidad, un tanto más complicado si queremos incluir el espín de los protones. Desde luego, ya no es correcto considerar a la molécula como un sistema de *dos* estados. Se la tendría que considerar como un sistema de *ocho* estados —hay cuatro arreglos posibles para el espín de cada uno de nuestros estados $|1\rangle$ y $|2\rangle$ — por lo que simplificamos demasiado al despreocupar los espines. Sin embargo, nuestras conclusiones finales son correctas.

Hemos encontrado que el estado de energía más baja —el único estado enlazado— de la molécula H_2 , tiene los dos electrones con sus espines opuestos. El momentum angular total de espín de los electrones es cero. Por otra parte, dos átomos de hidrógeno cercanos con espines paralelos —y, por consiguiente, con momentum angular total $\hbar$ — deben estar en un estado de energía más alto (no enlazado); los átomos se repelen entre sí. Hay una correlación interesante entre los espines y las energías. Nos proporciona otra ilustración de algo que mencionamos antes, que es que parece haber una energía de “interacción” entre dos espines, porque el caso de espines paralelos tiene una energía mayor que la del caso opuesto. En cierto sentido, podrían decir que los espines tratan de alcanzar un estado antiparalelo y, al hacerlo, tienen la capacidad de liberar energía —no porque haya una gran fuerza magnética, sino debido al principio de exclusión—.

Vimos en la sección 10-1 que el enlace de dos iones *diferentes* por medio de *un solo* electrón es probablemente muy débil. Esto *no* es cierto para el enlace con *dos* electrones. Supongamos que los dos protones en la figura 10-4 fueran remplazados por dos iones cualesquiera (con capas electrónicas internas cerradas y una sola carga iónica) y que las energías de enlace de un electrón en los dos iones fueran diferentes. Las energías de los estados $|1\rangle$ y $|2\rangle$ seguirían siendo iguales, porque en cada uno de estos estados tenemos un electrón enlazado a cada ion. En consecuencia, siempre tenemos la separación proporcional a A . El enlace con dos electrones es omnipresente —es el enlace de valencia más común—. El enlace químico comúnmente implica este juego de ir y venir de los dos electrones. Aunque dos átomos pueden estar enlazados por un electrón solamente, esto es bastante raro —porque requiere justamente las condiciones precisas—.

Finalmente, queremos mencionar que si la energía de atracción de un electrón hacia un núcleo es mucho mayor que hacia el otro, entonces lo que hemos dicho antes sobre ignorar otros estados posibles ya no es correcto. Supongamos que el núcleo *a* (o también puede ser un ion positivo) atrae más fuertemente a un electrón que el núcleo *b*. Puede ocurrir entonces que la energía total siga siendo bastante baja, incluso cuando ambos electrones estén con el núcleo *a*, y ninguno esté en el núcleo *b*. La fuerte atracción puede compensar con creces la repulsión mutua de los dos electrones. Si es así, el estado de energía más baja puede tener una gran amplitud para encontrar ambos electrones en *a* (formando un ion negativo) y una pequeña amplitud para encontrar algún electrón en *b*. El estado se parece a un ion negativo con un ion positivo. Esto es, en realidad,

lo que ocurre en una molécula “iónica” como NaCl. Pueden ver que todas las gradaciones entre enlace covalente y enlace iónico son posibles.

Pueden empezar a darse cuenta por qué la descripción cuántica es la mejor manera de comprender muchos de los fenómenos de la química.

10-4 La molécula de benceno

Los químicos han inventado bellos diagramas para representar moléculas orgánicas complicadas. Ahora vamos a estudiar una de las más interesantes —la molécula de benceno mostrada en la figura 10-6—. Contiene seis átomos de carbono y seis átomos de hidrógeno en un arreglo simétrico. Cada barra del diagrama presenta a un *par* de electrones con espines opuestos, bailando la danza del enlace covalente. Cada átomo de hidrógeno contribuye con un electrón y cada átomo de carbono contribuye con cuatro electrones para un total de 30 electrones involucrados. (Hay dos electrones más cerca del núcleo del carbono, que forman la primera capa, o capa K. Éstos no se muestran porque están enlazados tan fuertemente, que no intervienen en forma apreciable en el enlace covalente.) Así, cada barra en la figura representa un *enlace*, o par de electrones, y los enlaces dobles significan que hay *dos pares* de electrones entre pares alternados de átomos de carbono.

Hay un misterio sobre la molécula de benceno. Podemos calcular cuál sería la energía necesaria para formar este compuesto químico, porque los químicos han medido las energías de varios compuestos en que incluyen pedazos del anillo —por ejemplo, conocen la energía de un enlace doble estudiando el etileno, etc.—. Así, podemos calcular la energía total que sería de esperar para la molécula de benceno. Sin embargo, la energía real del anillo bencénico es mucho más baja de lo que obtenemos con ese cálculo; está enlazada más fuertemente de lo que podría esperarse a partir del denominado “sistema de doble enlace insaturado”. Generalmente, un sistema de enlaces dobles, el cual no forma un anillo como éste, es fácil de atacar químicamente, porque tiene una energía bastante alta —es fácil romper los enlaces dobles agregando otros hidrógenos—. Pero en el benceno, el anillo es muy estable y difícil de romper. En otras palabras, el benceno tiene una energía mucho menor de la calculada con el modelo de enlaces.

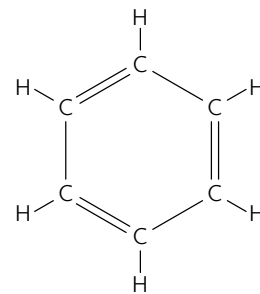


Figura 10-6. La molécula de benceno, C_6H_6 .

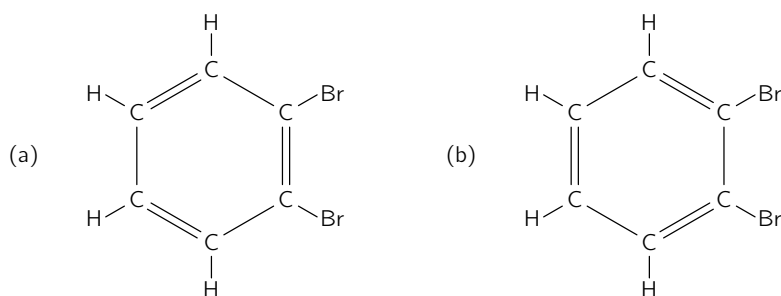


Figura 10-7. Dos posibilidades de orto-dibromobenceno. Los dos bromos podrían estar separados por un enlace simple o por un enlace doble.

Luego hay otro misterio. Supongan que remplazamos dos hidrógenos adyacentes por átomos de bromo para hacer orto-dibromobenceno. Hay dos maneras de hacer esto, como se muestra en la figura 10-7. Los bromos podrían estar en los extremos opuestos de un enlace doble, como en la parte (a) de la figura, o podrían estar en los extremos opuestos de un enlace simple, como en (b). Uno pensaría que el orto-dibromobenceno debe tener dos formas diferentes, pero no es así. Hay un solo compuesto de ese tipo.†

† Estamos simplificando un poco de más. Al principio, los químicos pensaban que tendría que haber *cuatro* formas de dibromobenceno: dos formas con los bromos sobre átomos de carbono adyacentes (orto-dibromobenceno), una tercera forma con los bromos sobre carbonos

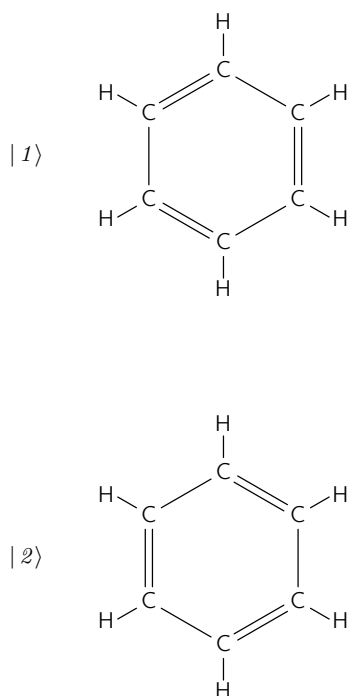


Figura 10-8. Un conjunto de estados base para la molécula de benceno.

Ahora queremos resolver estos misterios —y quizá ya hayan adivinado de qué modo: observando, naturalmente, que el “estado fundamental” del anillo bencénico es realmente un sistema de dos estados—. Podríamos imaginar que los enlaces del benceno pueden estar en cualquiera de las dos disposiciones mostradas en la figura 10-8. Ustedes dirán, “Pero en realidad son el mismo; debieran tener la misma energía”. De hecho, sí debieran. Y es por esa razón que se las debe analizar como un sistema de dos estados. Cada estado representa una configuración diferente de todo el conjunto de electrones, y hay cierta amplitud A para que todo el montón pase de un arreglo al otro —hay una posibilidad para que los electrones puedan cambiarse de una danza a la otra—.

Como hemos visto, esta posibilidad de saltar de un estado a otro da origen a un estado mixto, cuya energía es menor que la que se obtendría considerando de manera separada las dos imágenes de la figura 10-8. En su lugar, hay dos estados estacionarios —uno de energía mayor y otro de energía menor que el valor esperado—. En la realidad, el verdadero estado normal (de energía más baja) del benceno no es ninguna de las posibilidades mostradas en la figura 10-8, sino que tiene una amplitud $1/\sqrt{2}$ para estar en cada uno de los estados mostrados. Es el único estado que interviene en la química del benceno a temperaturas normales. Además, el estado mayor también existe; podemos decir que está ahí porque el benceno absorbe fuertemente la luz ultravioleta de frecuencia $\omega = (E_I - E_{II})/\hbar$. Recordarán que en el amoníaco, donde los objetos que cambiaban de lugar eran tres protones, la separación energética estaba en la región de las microondas. En el benceno, los objetos son electrones, y debido a que son mucho más ligeros, es más fácil que cambien de lugar, lo cual hace que el coeficiente A sea mucho mayor. El resultado es que la diferencia de energía es mucho mayor —alrededor de 1.5 eV, que es la energía de un fotón ultravioleta—.†

¿Qué pasa si sustituimos con bromo? Nuevamente las dos “posibilidades” (a) y (b) en la figura 10-7 representan las dos configuraciones electrónicas diferentes. La única diferencia es que los dos estados base con que comenzamos tendrán energías ligeramente diferentes. El estado estacionario de energía más baja seguirá siendo una combinación lineal de los dos estados, pero con amplitudes desiguales. La amplitud para el estado $|1\rangle$ podría valer algo así como $\sqrt{2/3}$, por ejemplo, mientras que la del estado $|2\rangle$ tendría una magnitud de $\sqrt{1/3}$. No podemos decir con seguridad sin tener más información, pero una vez que las dos energías H_{11} y H_{22} ya no son iguales, las amplitudes C_1 y C_2 ya no tienen magnitudes iguales. Esto significa, por supuesto, que una de las dos posibilidades en la figura es más probable que la otra, pero, como los electrones tienen movilidad suficiente, hay una amplitud para ambas. El otro estado tiene amplitudes diferentes (como $\sqrt{1/3}$ y $-\sqrt{2/3}$) pero está en una energía mayor. Hay un solo estado más bajo y no dos, como sugeriría la teoría ingenua de enlaces químicos fijos.

10-5 Colorantes

Les daremos un ejemplo más del fenómeno de dos estados en la química —esta vez en una escala molecular mayor—. Tiene que ver con la teoría de los colorantes. Muchos colorantes —en realidad, la mayoría de los colorantes artificiales— tienen una característica interesante; tienen un cierto tipo de simetría.

salteados (meta-dibromobenceno) y una cuarta con los bromos en posiciones opuestas (para-dibromobenceno). Sin embargo, sólo encontraron tres formas —sólo hay una forma de la orto-molécula—.

† Lo que hemos dicho nos puede conducir un poco a errores. La absorción de luz ultravioleta sería muy débil en el sistema de dos estados que hemos supuesto para el benceno, porque el elemento matricial del momento dipolar entre los dos estados es cero. [Los dos estados son simétricos eléctricamente, por lo que, en nuestra fórmula (9.55) para la probabilidad de una transición, el momento dipolar μ es cero y no hay absorción de luz.] Si éstos fueran los únicos estados, habría que demostrar la existencia del estado mayor de otra manera. Sin embargo, una teoría más completa del benceno, que inicia con más estados base (como los que tienen enlaces dobles adyacentes) muestra que los verdaderos estados estacionarios del benceno están ligeramente distorsionados respecto a los que hemos encontrado. Los momentos dipolares resultantes permiten que la transición mencionada en el texto ocurra por absorción de luz ultravioleta.

La figura 10-9 muestra un ion de un colorante llamado fucsina que tiene un color rojo púrpúreo. La molécula tiene tres estructuras cíclicas —dos de las cuales son anillos bencénicos—. La tercera no es exactamente igual a un anillo bencénico, porque sólo tiene dos enlaces dobles en el anillo. La figura muestra dos disposiciones igualmente satisfactorias, y podríamos imaginarnos que deben tener energías iguales. Pero hay una cierta amplitud para que todos los electrones puedan saltar de una configuración a la otra, cambiando la ubicación de la posición “sin llenar” al extremo opuesto. Como intervienen tantos electrones, la amplitud para cambiar de posición es algo más pequeña que en el caso del benceno, y la diferencia de energía entre los dos estados estacionarios es menor. No obstante, están los dos estados estacionarios habituales $|I\rangle$ y $|II\rangle$, que son la suma y la diferencia de los dos estados base mostrados en la figura. La separación energética entre $|I\rangle$ y $|II\rangle$ resulta ser igual a la energía de un fotón en la región visible. Si se hace brillar luz sobre la molécula, hay una absorción muy fuerte en una frecuencia, y aparece brillantemente coloreada. ¡Por eso es un colorante!

Otro rasgo interesante de dicha molécula de colorante es que, en los dos estados base mostrados, el centro de carga eléctrica está ubicado en lugares diferentes. Como resultado, la molécula debe ser afectada fuertemente por un campo eléctrico externo. Teníamos un efecto similar en la molécula de amoníaco. Es claro que podemos analizarlo empleando el mismo procedimiento matemático, siempre que conozcamos los números E_0 y A . Generalmente, éstos se obtienen de datos experimentales. Si realizamos mediciones con muchos colorantes, a veces es posible tener una idea de lo que ocurrirá con alguna molécula de colorante relacionada. A causa del gran cambio en la posición del centro de carga eléctrica, el valor de μ en la fórmula (9.55) es grande, y el material tiene una alta probabilidad para absorber luz de la frecuencia característica $2A/\hbar$. En consecuencia, no sólo está coloreado, sino que lo está muy intensamente —una pequeña cantidad de sustancia absorbe una gran cantidad de luz—.

La rapidez del cambio de posición —y, por lo tanto, A — es muy sensible a la estructura completa de la molécula. Al variar A , la separación energética, y con ella el color del colorante, se puede cambiar. Además, no es necesario que las moléculas sean perfectamente simétricas. Hemos visto que el mismo fenómeno básico existe con ligeras modificaciones, aunque haya una pequeña asimetría. Así, podemos modificar un poco los colores introduciendo ligeras asimetrías en las moléculas. Por ejemplo, otro colorante importante, el verde de malaquita, es muy similar a la fucsina, sólo que dos hidrógenos son remplazados por CH_3 . Es de un color diferente porque A se desplaza, y la rapidez del cambio de posición varía.

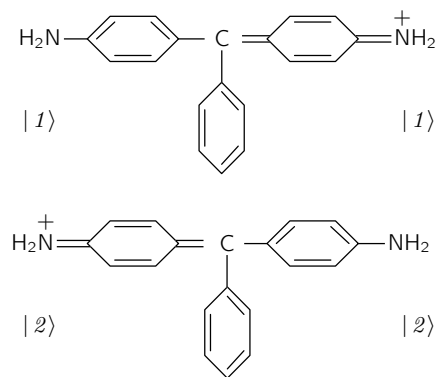


Figura 10-9. Dos estados base para la molécula del colorante magenta.

10-6 El hamiltoniano de una partícula con espín un medio en un campo magnético

Ahora queremos discutir un sistema de dos estados que involucra un objeto con espín un medio. Parte de lo que digamos ya ha sido tratado en capítulos anteriores, pero el tratarlo de nuevo puede ayudar a hacer más claros algunos de los detalles confusos. Podemos considerar a un electrón en reposo como un sistema de dos estados. Aunque en esta sección hablaremos de “un electrón”, lo que encontremos será válido para *cualquier* partícula con espín un medio. Supongan que escogemos como estados base a $|1\rangle$ y $|2\rangle$, los estados en los que la componente z del espín del electrón es $+\hbar/2$ y $-\hbar/2$.

Estos estados son, por supuesto, los mismos que hemos llamado (+) y (−) en capítulos anteriores. Para que la notación de este capítulo sea consistente, llamaremos al estado de espín “más” a $|1\rangle$ y estado de espín “menos” a $|2\rangle$ —donde “más” y “menos” se refieren al momentum angular en la dirección z —.

Cualquier estado posible ψ del electrón se puede describir como en la ecuación (10.1), al dar la amplitud C_1 para que el electrón esté en el estado $|1\rangle$, y la amplitud C_2 para que esté en el estado $|2\rangle$. Para tratar este problema, necesitaremos conocer el hamiltoniano para este sistema de dos estados —o sea, para un electrón en un campo magnético—. Comencemos con el caso especial de un campo magnético en la dirección z .

Supongan que el vector $\mathbf{B}$ tiene sólo a la componente z , B_z . Por la definición de los dos estados base (es decir, por tener espín paralelo o antiparalelo a $\mathbf{B}$) sabemos que éstos ya son estados estacionarios de energía definida en el campo magnético. El estado $|1\rangle$ corresponde a una energía[†] igual a $-\mu B_z$ y el estado $|2\rangle$ a $+\mu B_z$. El hamiltoniano tiene que ser muy simple en este caso, ya que C_1 , la amplitud para estar en el estado $|1\rangle$, no se ve afectada por C_2 , y viceversa:

$$\begin{aligned} i\hbar \frac{dC_1}{dt} &= E_1 C_1 = -\mu B_z C_1, \\ i\hbar \frac{dC_2}{dt} &= E_2 C_2 = +\mu B_z C_2. \end{aligned} \quad (10.17)$$

En este caso particular, el hamiltoniano es

$$\begin{aligned} H_{11} &= -\mu B_z, & H_{12} &= 0, \\ H_{21} &= 0, & H_{22} &= +\mu B_z. \end{aligned} \quad (10.18)$$

Entonces, sabemos cuál es el hamiltoniano para el campo magnético en la dirección z , y conocemos las energías de los estados estacionarios.

Supongan ahora que el campo *no* está en la dirección z . ¿Cuál es el hamiltoniano? ¿Cómo cambian los elementos de la matriz si el campo no está en la dirección z ? Vamos a suponer que hay algo como un principio de superposición para los términos del hamiltoniano. Más específicamente, queremos suponer que si dos campos magnéticos se superponen, los términos del hamiltoniano simplemente se suman —si conocemos los H_{ij} para un B_z puro y conocemos los H_{ij} para un B_x puro, entonces los H_{ij} para B_z y B_x juntos son simplemente la suma—. Esto es ciertamente válido si sólo consideramos campos en la dirección z —si duplicamos a B_z , entonces se duplican todos los H_{ij} —. Supongamos entonces que H es lineal en el campo $\mathbf{B}$. Esto es todo lo que necesitamos para poder encontrar los H_{ij} correspondientes a cualquier campo magnético.

Supongan que tenemos un campo constante $\mathbf{B}$. Podríamos haber elegido nuestro eje z en su dirección, y habríamos encontrado dos estados estacionarios con energías $\mp\mu B$. Sólo por cambiar la dirección de nuestros ejes no vamos a cambiar la *física*. Nuestra *descripción* de los estados estacionarios será diferente, pero sus energías *seguirán siendo* $\mp\mu B$; es decir,

$$E_I = -\mu \sqrt{B_x^2 + B_y^2 + B_z^2} \quad (10.19)$$

y

$$E_{II} = +\mu \sqrt{B_x^2 + B_y^2 + B_z^2}.$$

El resto del juego es fácil. Aquí tenemos las expresiones para las energías. Queremos un hamiltoniano que sea lineal en B_x , B_y y B_z , y que dé estas energías al usarlo en nuestra fórmula general (10.3). El problema: encontrar el hamiltoniano. Primero observen que la separación de energía es simétrica, tenemos un valor promedio de cero. Examinando la ecuación (10.3), vemos directamente que eso requiere

$$H_{22} = -H_{11}.$$

(Observen que esto concuerda con lo que ya sabemos cuando ambos B_x y B_y son cero; en ese caso $H_{11} = -\mu B_z$, y $H_{22} = \mu B_z$.) Ahora bien, si igualamos las energías de la ecuación (10.3) con lo que sabemos de la ecuación (10.19), tenemos

$$\left(\frac{H_{11} - H_{22}}{2} \right)^2 + |H_{12}|^2 = \mu^2 (B_x^2 + B_y^2 + B_z^2). \quad (10.20)$$

[†] Estamos tomando la energía en reposo $m_0 c^2$ como nuestro “cero” de energía y tratando al momento magnético μ del electrón como un número *negativo*, ya que apunta en dirección opuesta al espín.

(También hemos usado la propiedad $H_{21} = H_{12}^*$, por lo que $H_{12}H_{21}$ se puede escribir como $|H_{12}|^2$.) De nuevo, en el caso especial de un campo en la dirección z , se obtiene

$$\mu^2 B_z^2 + |H_{12}|^2 = \mu^2 B_z^2.$$

Es claro que $|H_{12}|$ debe ser cero en este caso especial, lo cual significa que H_{12} no puede contener ningún término en B_z . (Recuerden que hemos dicho que todos los términos deben ser lineales en B_x , B_y y B_z .)

Hasta el momento, entonces, hemos descubierto que H_{11} y H_{22} tienen términos en B_z , mientras que H_{12} y H_{21} no. Podemos hacer la siguiente hipótesis para satisfacer la ecuación (10.20)

$$\begin{aligned} H_{11} &= -\mu B_z, \\ H_{22} &= \mu B_z, \end{aligned} \quad (10.21)$$

y

$$|H_{12}|^2 = \mu^2 (B_x^2 + B_y^2).$$

¡Y resulta que ésa es la única manera posible!

“Espere” —ustedes dirán— “ H_{12} no es lineal en B ; la ecuación (10.21) da $H_{12} = \mu \sqrt{B_x^2 + B_y^2}$ ”. No necesariamente. Hay otra posibilidad que *es* lineal, es decir,

$$H_{12} = \mu(B_x + iB_y).$$

En realidad, hay varias posibilidades de este tipo —en forma completamente general, podríamos escribir

$$H_{12} = \mu(B_x \pm iB_y)e^{i\delta},$$

donde δ es una fase arbitraria—. ¿Qué signo y qué fase debemos usar? Resulta que pueden elegir cualquier signo, y cualquier fase que quieran, y los resultados físicos serán siempre los mismos. Por lo tanto, la elección es cuestión de convención. Los que nos han precedido escogieron usar el signo menos y tomar $e^{i\delta} = -1$. Podemos seguir la convención y escribir

$$H_{12} = -\mu(B_x - iB_y), \quad H_{21} = -\mu(B_x + iB_y).$$

(Incidentalmente, estas convenciones están relacionadas, y son consistentes, con algunas elecciones arbitrarias que hicimos en el capítulo 6.)

El hamiltoniano completo para un electrón en un campo magnético arbitrario es, entonces,

$$\begin{aligned} H_{11} &= -\mu B_z, & H_{12} &= -\mu(B_x - iB_y), \\ H_{21} &= -\mu(B_x + iB_y), & H_{22} &= +\mu B_z. \end{aligned} \quad (10.22)$$

Y las ecuaciones para las amplitudes C_1 y C_2 son

$$\begin{aligned} i\hbar \frac{dC_1}{dt} &= -\mu[B_z C_1 + (B_x - iB_y)C_2], \\ i\hbar \frac{dC_2}{dt} &= -\mu[(B_x + iB_y)C_1 - B_z C_2]. \end{aligned} \quad (10.23)$$

De este modo, hemos descubierto las “ecuaciones de movimiento de los estados con espín” de un electrón en un campo magnético. Las adivinamos, por decirlo así, haciendo algunos razonamientos físicos, pero la prueba real de cualquier hamiltoniano es que debe dar predicciones que estén de acuerdo con los resultados experimentales. Según todas las pruebas realizadas, estas ecuaciones son correctas. En realidad, aunque hicimos nuestros razonamientos sólo para campos constantes, el hamiltoniano que hemos escrito también es correcto para campos magnéticos que varían con el tiempo. Por ello, podemos usar ahora la ecuación (10.23) para examinar toda clase de problemas interesantes.

10-7 El electrón con espín en un campo magnético

Ejemplo número uno: comenzamos con un campo constante en la dirección z . Hay sólo dos estados estacionarios con energías $\mp \mu B_z$. Supongan que agregamos un campo pequeño en la dirección x . Entonces, las ecuaciones se parecen a nuestro repetido problema de dos estados. Surge, una vez más, el problema del cambio de lugar, y los niveles de energía se separan un poco. Ahora, hagamos que la componente x del campo varíe con el tiempo —digamos que como $\cos \omega t$ —. Las ecuaciones son entonces las mismas que teníamos cuando aplicábamos un campo eléctrico oscilante a la molécula de amoníaco en el capítulo 9. Pueden calcular los detalles en la misma forma. Obtendrán el resultado de que el campo oscilante provoca transiciones del estado $+z$ al estado $-z$ —y viceversa— cuando el campo horizontal oscila cerca de la frecuencia de resonancia $\omega_0 = 2\mu B_z/\hbar$. *Esto da la teoría cuántica de los fenómenos de resonancia magnética que describimos en el capítulo 35 del volumen II (véase el apéndice).*

También es posible hacer un máser que utilice un sistema de espín un medio. Se usa un aparato de Stern-Gerlach para producir un haz de partículas polarizadas en, por ejemplo, la dirección $+z$, las cuales se envían dentro de una cavidad con un campo magnético constante. Los campos oscilantes que hay dentro de la cavidad se pueden acoplar con el momento magnético e inducir transiciones que dan energía a la cavidad.

Consideremos ahora la siguiente pregunta. Supongan que tenemos un campo magnético $\mathbf{B}$, el cual apunta en la dirección cuyo ángulo polar es θ y ángulo azimutal es ϕ , como en la figura 10-10. Supongan, además, que hay un electrón que ha sido preparado con su espín apuntando en la dirección de este campo. ¿Cuáles son las amplitudes C_1 y C_2 para dicho electrón? En otras palabras, llamando al estado del electrón $|\psi\rangle$, queremos escribir

$$|\psi\rangle = |1\rangle C_1 + |2\rangle C_2,$$

donde C_1 y C_2 son

$$C_1 = \langle 1 | \psi \rangle, \quad C_2 = \langle 2 | \psi \rangle,$$

donde por $|1\rangle$ y $|2\rangle$ entendemos lo mismo que solíamos llamar $|+\rangle$ y $|-\rangle$ (referidos al eje z que hemos elegido).

La respuesta a esta pregunta también está en las ecuaciones generales para los sistemas de dos estados. En primer lugar, sabemos que como el espín del electrón es paralelo a $\mathbf{B}$, está en un estado estacionario de energía $E_I = -\mu B$. En consecuencia, tanto C_1 como C_2 deben variar como $e^{-iE_I t/\hbar}$, como en la ecuación (9.18); y sus coeficientes a_1 y a_2 están dados por (10.5), o sea,

$$\frac{a_1}{a_2} = \frac{H_{12}}{E_I - H_{11}}. \quad (10.24)$$

Una condición adicional es que a_1 y a_2 deben estar normalizados de modo que $|a_1|^2 + |a_2|^2 = 1$. Podemos tomar H_{11} y H_{12} de (10.22) usando

$$B_z = B \cos \theta, \quad B_x = B \sin \theta \cos \phi, \quad B_y = B \sin \theta \sin \phi.$$

Tenemos entonces,

$$\begin{aligned} H_{11} &= -\mu B \cos \theta, \\ H_{12} &= -\mu B \sin \theta (\cos \phi - i \sin \phi). \end{aligned} \quad (10.25)$$

El último factor en la segunda ecuación es, además, $e^{-i\phi}$, por lo que es más simple escribir

$$H_{12} = -\mu B \sin \theta e^{-i\phi}. \quad (10.26)$$

Usando estos elementos matriciales en la ecuación (10.24) —y cancelando $-\mu B$ del numerador y el denominador— encontramos

$$\frac{a_1}{a_2} = \frac{\sin \theta e^{-i\phi}}{1 - \cos \theta}. \quad (10.27)$$

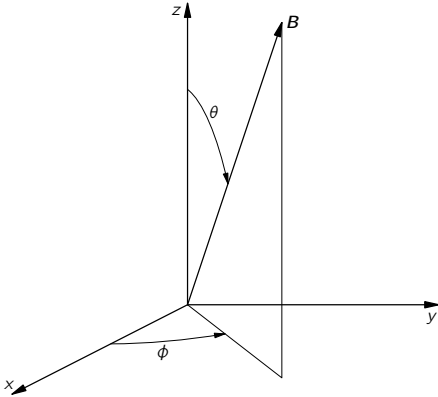


Figura 10-10. La dirección de $\mathbf{B}$ está definida por el ángulo polar θ y el ángulo azimutal ϕ .

Con este cociente, y la condición de normalización, podemos encontrar ambos a_1 y a_2 . No es difícil, pero podemos tomar un atajo con un pequeño truco. Observen que $1 - \cos \theta = 2 \sin^2(\theta/2)$, y que $\sin \theta = 2 \sin(\theta/2) \cos(\theta/2)$. Entonces, la ecuación (10.27) es equivalente a

$$\frac{a_1}{a_2} = \frac{\cos \frac{\theta}{2} e^{-i\phi}}{\sin \frac{\theta}{2}}. \quad (10.28)$$

Por lo tanto, una respuesta posible es

$$a_1 = \cos \frac{\theta}{2} e^{-i\phi}, \quad a_2 = \sin \frac{\theta}{2}, \quad (10.29)$$

ya que satisface (10.28) y también cumple con

$$|a_1|^2 + |a_2|^2 = 1.$$

Como ustedes saben, multiplicar a ambos a_1 y a_2 por un factor de fase arbitrario no cambia nada. Generalmente, la gente prefiere hacer las ecuaciones (10.29) más simétricas al multiplicar ambas por $e^{i\phi/2}$. De ahí que la forma que se usa comúnmente sea

$$a_1 = \cos \frac{\theta}{2} e^{-i\phi/2}, \quad a_2 = \sin \frac{\theta}{2} e^{+i\phi/2}, \quad (10.30)$$

y ésta es la respuesta a nuestra pregunta. Los números a_1 y a_2 son las amplitudes para encontrar a un electrón con su espín hacia arriba, o hacia abajo, en la dirección del eje z , cuando sabemos que su espín está en la dirección del eje en θ y ϕ . (Las amplitudes C_1 y C_2 son simplemente a_1 y a_2 por $e^{-iE_I t/\hbar}$.)

Ahora observen una cosa interesante. La intensidad B del campo magnético no aparece por ningún lado en (10.30). El resultado es evidentemente el mismo en el límite cuando B tiende a cero. Esto significa que hemos dado una respuesta *general* a la pregunta de cómo representar una partícula cuyo espín está en la dirección de un eje arbitrario. Las amplitudes que aparecen en (10.30) son las amplitudes para la proyección de partículas con espín un medio, que corresponden a las amplitudes para la proyección que dimos en el capítulo 5 [ecuaciones (5.38)] para partículas con espín uno. Ahora podemos encontrar las amplitudes para que haces filtrados de partículas con espín un medio atraviesen cualquier filtro de Stern-Gerlach.

Con $|+z\rangle$ representamos a un estado con espín hacia arriba en la dirección del eje z , y con $|-z\rangle$ representamos al estado con espín hacia abajo. Si $|+z'\rangle$ representa a un estado con espín hacia arriba en la dirección de un eje z' que forma los ángulos polares θ y ϕ con el eje z , entonces, en la notación del capítulo 5, tenemos

$$|+z|+z'\rangle = \cos \frac{\theta}{2} e^{-i\phi/2}, \quad \langle -z|+z'\rangle = \sin \frac{\theta}{2} e^{+i\phi/2}. \quad (10.31)$$

Estos resultados son equivalentes a los del capítulo 6, ecuación (6.36), que encontramos con argumentos puramente geométricos. (Así que si decidieron omitir el capítulo 6, ahora tienen los resultados esenciales de todas maneras.)

Como ejemplo final, examinemos nuevamente uno que ya hemos mencionado varias veces. Supongan que consideramos el siguiente problema. Comenzamos con un electrón cuyo espín está en cierta dirección, después encendemos un campo magnético en la dirección z durante 25 minutos, y luego lo apagamos. ¿Cuál es el estado final? De nuevo, representemos al estado por la combinación lineal $|\psi\rangle = |1\rangle C_1 + |2\rangle C_2$. En este problema, sin embargo, los estados de energía definida son también nuestros estados base $|1\rangle$ y $|2\rangle$. Por lo tanto, C_1 y C_2 sólo varían en la fase. Sabemos que

$$C_1(t) = C_1(0)e^{-iE_I t/\hbar} = C_1(0)e^{+i\mu B t/\hbar},$$

y

$$C_2(t) = C_2(0)e^{-iE_{II} t/\hbar} = C_2(0)e^{-i\mu B t/\hbar}.$$

Ahora bien, dijimos al principio que el espín del electrón estaba en una dirección determinada. Eso significa que inicialmente C_1 y C_2 son dos números dados por las ecuaciones (10.30). Después de esperar un tiempo T , los nuevos C_1 y C_2 son los mismos dos números multiplicados por $e^{i\mu B_z T/\hbar}$ y $e^{-i\mu B_z T/\hbar}$, respectivamente. ¿Qué estado es ése? Es fácil. Es exactamente el mismo que si el ángulo ϕ se hubiera cambiado restándole $2\mu B_z T/\hbar$ y el ángulo θ no se hubiera cambiado. Eso significa que al final del tiempo T , el estado $|\psi\rangle$ representa a un electrón alineado en una dirección que difiere de la dirección original sólo por una *rotación* alrededor del eje z en un ángulo $\Delta\phi = 2\mu B_z T/\hbar$. Como este ángulo es proporcional a T , también podemos decir que la dirección del espín *precesa* con velocidad angular $2\mu B_z/\hbar$ alrededor del eje z . Ya hemos discutido varias veces antes este resultado en una forma menos completa y rigurosa. Ahora hemos obtenido una descripción cuántica completa y precisa de la precesión de los imanes atómicos.

Es interesante que las ideas matemáticas que acabamos de desarrollar para el electrón con espín en un campo magnético, se pueden aplicar a *cualquier* sistema de dos estados. Esto significa que, haciendo una *analogía* matemática con el electrón con espín, *cualquier problema* sobre sistemas de dos estados se puede resolver utilizando geometría pura. El funcionamiento es el siguiente. Primero cambian el cero de energía de modo que $(H_{11} + H_{22})$ sea igual a cero, por lo que $H_{11} = -H_{22}$. Entonces, cualquier problema de dos estados es *formalmente* igual al del electrón en un campo magnético. Todo lo que tienen que hacer es *identificar* $-\mu B_z$ con H_{11} y $-\mu(B_x - iB_y)$ con H_{12} . Sin importar cuál sea la situación física original —una molécula de amoníaco, o lo que sea— la pueden traducir al correspondiente problema de electrones. Así entonces, si podemos resolver el problema de electrones *en general*, hemos resuelto *todos* los problemas de dos estados.

¡Y tenemos la solución general para el electrón! Supongan que tienen un estado para empezar, el cual tiene el espín “hacia arriba” en alguna dirección, y que tienen un campo magnético $\mathbf{B}$ que apunta en alguna otra dirección. Sólo rotan la dirección del espín alrededor del eje de $\mathbf{B}$ con velocidad angular *vectorial* $\boldsymbol{\omega}(t)$ igual a una constante por el vector $\mathbf{B}$ (es decir, $\boldsymbol{\omega} = 2\mu\mathbf{B}/\hbar$). A medida que $\mathbf{B}$ varía con el tiempo, continúan moviendo el eje de rotación para mantenerlo paralelo a $\mathbf{B}$ y siguen cambiando la velocidad de rotación para que siempre sea proporcional a la intensidad de $\mathbf{B}$. Véase la figura 10-11. Si siguen haciendo esto, terminarán con una cierta orientación final para el eje del espín, y las amplitudes C_1 y C_2 estarán dadas precisamente por las proyecciones —usando (10.30)— sobre su sistema de coordenadas. Ya ven, es simplemente un problema geométrico de no perder de vista en dónde terminan después de rotar. Aunque es fácil ver todo lo que está implicado, este problema geométrico (de encontrar el resultado de una rotación con un vector velocidad angular variable) no es fácil de resolver explícitamente en el caso general. De todas maneras, vemos cuál es, *en principio*, la solución general de cualquier problema de dos estados. En el próximo capítulo, examinaremos un poco más las técnicas matemáticas para tratar el caso importante de una partícula con espín un medio —y, en consecuencia, para tratar sistemas de dos estados en general—.

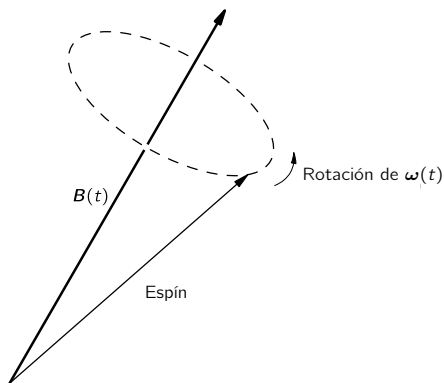


Figura 10-11. La dirección del espín de un electrón en un campo magnético variable $\mathbf{B}(t)$ precesa a la frecuencia $\boldsymbol{\omega}(t)$ alrededor de un eje paralelo a $\mathbf{B}$.

Más sistemas de dos estados

11-1 Las matrices de espín de Pauli

Continuamos nuestra discusión sobre los sistemas de dos estados. Al final del capítulo anterior estuvimos hablando de una partícula de espín un medio en un campo magnético. Describimos el estado de espín dando la amplitud C_1 para que la componente z del momento angular del espín sea $+\hbar/2$, y la amplitud C_2 para que sea $-\hbar/2$. En capítulos anteriores llamamos a estos estados base $|+\rangle$ y $|-\rangle$. Usaremos ahora esta notación, aunque ocasionalmente nos sea conveniente usar $|+\rangle$ o $|1\rangle$, y $|-\rangle$ o $|2\rangle$, de manera alternativa.

Vimos, en el capítulo anterior, que cuando una partícula de espín un medio con un momento magnético μ se encuentra en un campo magnético $\mathbf{B} = (B_x, B_y, B_z)$, las amplitudes C_+ ($= C_1$) y C_- ($= C_2$) están relacionadas mediante las siguientes ecuaciones diferenciales:

$$\begin{aligned} i\hbar \frac{dC_+}{dt} &= -\mu[B_z C_+ + (B_x - iB_y)C_-], \\ i\hbar \frac{dC_-}{dt} &= -\mu[(B_x + iB_y)C_+ - B_z C_-]. \end{aligned} \quad (11.1)$$

En otras palabras, la matriz hamiltoniana H_{ij} es

$$\begin{aligned} H_{11} &= -\mu B_z, & H_{12} &= -\mu(B_x - iB_y), \\ H_{21} &= -\mu(B_x + iB_y), & H_{22} &= +\mu B_z. \end{aligned} \quad (11.2)$$

Y las ecuaciones (11.1) son, naturalmente, lo mismo que

$$i\hbar \frac{dC_i}{dt} = \sum_j H_{ij} C_j, \quad (11.3)$$

donde i y j toman los valores $+$ y $-$ (o 1 y 2).

El sistema de dos estados del espín del electrón es tan importante, que es muy útil tener un modo más claro para escribir las cosas. Ahora, haremos una pequeña digresión matemática para mostrarles cómo se suelen escribir las ecuaciones de un sistema de dos estados. Se hace así: primero, observen que cada término del hamiltoniano es proporcional a μ y a alguna componente de $\mathbf{B}$; entonces —de un modo puramente formal— podemos escribir

$$H_{ij} = -\mu[\sigma_{ij}^x B_x + \sigma_{ij}^y B_y + \sigma_{ij}^z B_z]. \quad (11.4)$$

Aquí no hay nada nuevo, desde el punto de vista físico; esta ecuación sólo significa que los coeficientes σ_{ij}^x , σ_{ij}^y y σ_{ij}^z —hay $4 \times 3 = 12$ de ellos— se pueden deducir para que (11.4) sea idéntica a (11.2).

Veamos cómo tienen que ser. Empezamos con B_z . Ya que B_z sólo aparece en H_{11} y H_{22} , todo estará bien si

$$\begin{aligned} \sigma_{11}^z &= 1, & \sigma_{12}^z &= 0, \\ \sigma_{21}^z &= 0, & \sigma_{22}^z &= -1. \end{aligned}$$

Con frecuencia, escribimos a la matriz H_{ij} en forma de una pequeña tabla, así:

$$H_{ij} = i\hbar \begin{pmatrix} H_{11} & H_{12} \\ H_{21} & H_{22} \end{pmatrix}.$$

11-1 Las matrices de espín de Pauli

11-2 Las matrices de espín como operadores

11-3 La solución de las ecuaciones de dos estados

11-4 Los estados de polarización del fotón

11-5 El mesón K neutro†

11-6 Generalización a sistemas de N estados

Revisar

Volumen I, capítulo 33,
Polarización

† Esta sección se debería omitir al leer el libro por primera vez. Es más avanzada de lo que corresponde a un primer curso.

Para el hamiltoniano de una partícula con espín un medio en el campo magnético B_z , esto es lo mismo que

$$H_{ij} = {}^{i\downarrow j\rightarrow} \begin{pmatrix} -\mu B_z & -\mu(B_x - iB_y) \\ -\mu(B_x + iB_y) & +\mu B_z \end{pmatrix}.$$

Del mismo modo, podemos escribir los coeficientes σ_{ij}^z como la matriz

$$\sigma_{ij}^z = {}^{i\downarrow j\rightarrow} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \quad (11.5)$$

Trabajando con los coeficientes de B_x , obtenemos que los términos de σ_x tienen que ser

$$\begin{aligned} \sigma_{11}^x &= 0, & \sigma_{12}^x &= 1, \\ \sigma_{21}^x &= 1, & \sigma_{22}^x &= 0. \end{aligned}$$

O, de manera abreviada,

$$\sigma_{ij}^x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}. \quad (11.6)$$

Finalmente, si consideramos a B_y , tenemos

$$\begin{aligned} \sigma_{11}^y &= 0, & \sigma_{12}^y &= -i, \\ \sigma_{21}^y &= i, & \sigma_{22}^y &= 0; \end{aligned}$$

o

$$\sigma_{ij}^y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}. \quad (11.7)$$

Tabla 11-1

Las matrices de espín de Pauli

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

Con estas tres matrices sigma, las ecuaciones (11.2) y (11.4) son idénticas. Para dejar espacio para los subíndices i y j , hemos mostrado que σ va con cada una de las componentes de $\mathbf{B}$ colocando como superíndices x , y y z . Sin embargo, por lo general i y j se omiten —es fácil imaginar que se encuentran ahí— y se usa x , y y z como subíndices. De esta forma, la ecuación (11.4) se escribe

$$H = -\mu[\sigma_x B_x + \sigma_y B_y + \sigma_z B_z]. \quad (11.8)$$

Puesto que las matrices sigma son tan importantes —los profesionales las usan continuamente— las hemos colocado juntas en la tabla 11-1. (Quien vaya a trabajar en física cuántica realmente tiene que memorizarlas.) También se les llama las *matrices de espín de Pauli*, en honor del físico que las inventó.

En la tabla hemos incluido una matriz de 2×2 más, la cual se necesita si se quiere ser capaz de tratar un sistema que tenga dos estados de espín de la misma energía, o si se quiere escoger una energía cero diferente. Para estos casos, tenemos que sumar $E_0 C_+$ a la primera ecuación en (11.1) y $E_0 C_-$ a la segunda. Podemos incluir esto en la nueva notación si definimos la *matriz unitaria* “1” como δ_{ij} ,

$$1 = \delta_{ij} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad (11.9)$$

y escribimos la ecuación (11.8) en la forma

$$H = E_0 \delta_{ij} - \mu(\sigma_x B_x + \sigma_y B_y + \sigma_z B_z). \quad (11.10)$$

Usualmente, se *entiende* que cualquier constante, como E_0 , tiene que ser automáticamente multiplicada por la matriz unidad; y de ahí que se escriba simplemente

$$H = E_0 - \mu(\sigma_x B_x + \sigma_y B_y + \sigma_z B_z). \quad (11.11)$$

Una razón por la que las matrices de espín son útiles es que *cualquier* matriz de 2×2 se puede expresar en términos de ellas. Cualquier matriz que ustedes escriban tiene cuatro números, por ejemplo,

$$M = \begin{pmatrix} a & b \\ c & d \end{pmatrix}.$$

Y siempre se puede escribir como una combinación lineal de cuatro matrices. Por ejemplo,

$$M = a \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} + b \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} + c \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} + d \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}.$$

Hay muchas formas de hacerlo, pero una muy especial es decir que M es una cierta cantidad por σ_x , más otra cierta cantidad por σ_y , etc., en esta forma:

$$M = \alpha 1 + \beta \sigma_x + \gamma \sigma_y + \delta \sigma_z,$$

donde las “cantidades” α , β , γ y δ pueden ser, en general, números complejos.

Como cualquier matriz de 2×2 se puede representar en términos de las matrices unidad y las matrices sigma, tenemos todo lo que alguna vez podamos necesitar para *cualquier* sistema de dos estados. Sin importar cuál sea el sistema de dos estados —la molécula de amoníaco, el colorante fucsina, o cualquier otra—, la ecuación hamiltoniana se puede escribir en términos de las sigmas. Aunque parece que las sigmas tienen un significado geométrico en la situación física de un electrón en un campo magnético, también pueden ser vistas sólo como matrices muy útiles, que se pueden usar para cualquier problema de dos estados.

Por ejemplo, en una manera de ver las cosas, un protón y un neutrón se pueden considerar como la misma partícula en cualquiera de los dos estados. Decimos que el *nucleón* (protón o neutrón) es un sistema de dos estados —en este caso, dos estados respecto a su carga—. Al considerarlo así, el estado $|1\rangle$ puede representar al protón y el estado $|2\rangle$ al neutrón. Algunos dicen que el nucleón tiene dos estados de “espín isotópico”.

Ya que vamos a usar las matrices sigma como la “aritmética” de la mecánica cuántica para los sistemas de dos estados, revisemos rápidamente las convenciones del álgebra de matrices. Por la “suma” de cualesquiera dos o más matrices, entendemos simplemente lo que era obvio en la ecuación (11.4). En general, si “sumamos” dos matrices A y B , la “suma” C significa que cada término C_{ij} está dado por

$$C_{ij} = A_{ij} + B_{ij}.$$

Cada término de C es la suma de los términos de la misma posición de A y B .

Como en la sección 5-6 ya introdujimos la idea de una matriz “producto”. La misma idea será útil al tratar con las matrices sigma. En general, el “producto” de dos matrices A y B (en este orden) se define como la matriz C cuyos elementos son

$$C_{ij} = \sum_k A_{ik} B_{kj}. \quad (11.12)$$

Es la suma de productos de términos, tomados en pares del renglón i -ésimo de A y de la columna j -ésima de B . Si las matrices están escritas en la forma de tabla, como en la figura 11-1, hay un buen “sistema” para obtener los términos de la matriz producto. Supongan que está calculando C_{23} . Recorren con su índice izquierdo el *segundo renglón* de A y con el índice derecho *hacia abajo* de la *tercera columna* de B , multiplicando cada par y sumándolo a medida que avanzan. En la figura hemos tratado de indicar cómo hacerlo.

Es, claro está, particularmente simple para matrices de 2×2 . Por ejemplo, si multiplicamos σ_x por σ_x , obtenemos

$$\sigma_x^2 = \sigma_x \cdot \sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \cdot \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix},$$

que es justamente la matriz unidad 1. Tomando otro ejemplo, calculemos $\sigma_x \sigma_y$:

$$\sigma_x \sigma_y = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \cdot \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} = \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix}.$$

Refiriéndonos a la tabla 11-1, ven que el producto es sólo i veces la matriz σ_z . (Recuerden que un número por una matriz consiste en multiplicar cada término

$$\begin{pmatrix} A_{11} & A_{12} & A_{13} & A_{14} \\ A_{21} & A_{22} & A_{23} & A_{24} \\ A_{31} & A_{32} & A_{33} & A_{34} \\ A_{41} & A_{42} & A_{43} & A_{44} \end{pmatrix} \cdot \begin{pmatrix} B_{11} & B_{12} & B_{13} & B_{14} \\ B_{21} & B_{22} & B_{23} & B_{24} \\ B_{31} & B_{32} & B_{33} & B_{34} \\ B_{41} & B_{42} & B_{43} & B_{44} \end{pmatrix} = \begin{pmatrix} C_{11} & C_{12} & C_{13} & C_{14} \\ C_{21} & C_{22} & C_{23} & C_{24} \\ C_{31} & C_{32} & C_{33} & C_{34} \\ C_{41} & C_{42} & C_{43} & C_{44} \end{pmatrix}$$

$$C_{ij} = \sum_k A_{ik} B_{kj}$$

Ejemplo: $C_{23} = A_{21}B_{13} + A_{22}B_{23} + A_{23}B_{33} + A_{24}B_{43}$

Figura 11-1. Multiplicación de dos matrices.

Tabla 11-2

Productos de las matrices de espín

$\sigma_x^2 = 1$
$\sigma_y^2 = 1$
$\sigma_z^2 = 1$
$\sigma_x \sigma_y = -\sigma_y \sigma_x = i\sigma_z$
$\sigma_y \sigma_z = -\sigma_z \sigma_y = i\sigma_x$
$\sigma_z \sigma_x = -\sigma_x \sigma_z = i\sigma_y$

de la matriz por él.) Como los productos de las sigmas, tomadas dos a la vez, son importantes —además de bastante divertidos— los hemos colocado en la tabla 11-2. Los pueden calcular como hemos hecho con σ_x^2 y $\sigma_x \sigma_y$.

Hay otro punto muy importante e interesante acerca de las matrices σ . Podemos imaginar, si queremos, que las matrices σ_x , σ_y y σ_z son análogas a las tres componentes de un vector —se le llama a veces el “vector sigma” y se escribe $\boldsymbol{\sigma}$ —. Es realmente un “vector matricial” o una “matriz vectorial”. Son tres matrices diferentes —cada matriz asociada a uno de los tres ejes x , y y z —. Con ese vector matricial, podemos escribir el hamiltoniano del sistema en una forma que sirve para cualquier sistema de coordenadas:

$$H = -\mu \boldsymbol{\sigma} \cdot \mathbf{B}. \quad (11.13)$$

Aunque hemos escrito nuestras tres matrices en la representación donde “hacia arriba” y “hacia abajo” están en la dirección z —de modo que σ_z tiene una simplicidad particular— podríamos saber cómo serían las matrices en cualquier otra representación. Aunque requiere mucha álgebra, pueden demostrar que cambian entre sí como las componentes de un vector. (No nos preocuparemos de probarlo por ahora. Ustedes pueden comprobarlo si quieren.) Pueden usar a $\boldsymbol{\sigma}$ en diferentes sistemas de coordenadas como si fuese un vector.

Recuerdan que H está relacionada con la energía en la mecánica cuántica. Es, de hecho, exactamente igual a la energía en el caso simple en que hay sólo un estado. Aun para sistemas de dos estados del espín del electrón, cuando escribimos el hamiltoniano como en la ecuación (11.13), se parece mucho a la fórmula clásica para la energía de un pequeño imán con momento magnético $\boldsymbol{\mu}$ en un campo magnético $\mathbf{B}$. De manera clásica, diríamos

$$U = -\boldsymbol{\mu} \cdot \mathbf{B}, \quad (11.14)$$

donde $\boldsymbol{\mu}$ es una propiedad del objeto y $\mathbf{B}$ es un campo externo. Podemos imaginar que la ecuación (11.14) se puede convertir en (11.13) si remplazamos a la energía clásica por el hamiltoniano y la $\boldsymbol{\mu}$ clásica por la matriz $\mu \boldsymbol{\sigma}$. Entonces, después de esta sustitución puramente formal, interpretamos el resultado como una ecuación matricial. Se dice algunas veces que a cada cantidad en física clásica le corresponde una matriz en mecánica cuántica. En realidad es más correcto decir que la matriz hamiltoniana corresponde a la energía, y que si otra cantidad se puede definir a través de la energía entonces tiene una matriz correspondiente.

Por ejemplo, el momento magnético se puede definir a través de la energía, diciendo que la energía en un campo externo $\mathbf{B}$ es $-\boldsymbol{\mu} \cdot \mathbf{B}$. Esto define al vector de momento magnético $\boldsymbol{\mu}$. Entonces observamos a la fórmula del hamiltoniano para objeto real (cuántico) en un campo magnético, y tratamos de identificar cuáles son las matrices que corresponden a las cantidades de la fórmula clásica. Éste es el truco por el cual algunas veces las cantidades clásicas tienen su contraparte cuántica.

Pueden tratar de entender, si quieren, cómo es que un vector clásico es igual a la matriz $\mu\sigma$, y quizá descubrirán algo —pero no se rompan la cabeza—. Ésa no es la idea —*no son iguales*—. La mecánica cuántica es una teoría diferente para representar al mundo. Solamente sucede que hay ciertas correspondencias, que apenas son algo más que artificios nemotécnicos —los cuales nos ayudan a recordar—. Esto es, ustedes recuerdan la ecuación (11.14) de cuando aprendieron en física clásica; entonces, si recuerdan la correspondencia $\mu \rightarrow \mu\sigma$, tienen un medio para recordar la ecuación (11.13). Por supuesto, la naturaleza conoce la mecánica cuántica, y la mecánica clásica es sólo una aproximación; así entonces, no es ningún misterio que en la mecánica clásica haya alguna sombra de las leyes cuánticas —que son las que verdaderamente se encuentran en la base de todo—. No es posible reconstruir de manera directa al objeto original a partir de la sombra, pero la sombra sí les ayuda a recordar a qué se parece el objeto. La ecuación (11.13) es la verdad, y la (11.14) es la sombra. Ya que aprendimos primero física clásica, quisiéramos ser capaces de obtener la fórmula cuántica de ella; pero no hay un modo infalible para hacerlo. Siempre tenemos que volver al mundo real y descubrir las ecuaciones cuánticas correctas. Cuando se parecen un poco a la física clásica, somos afortunados.

Si creen que las advertencias anteriores son repetitivas, y les parecen verdades demasiado elaboradas, y evidentes por sí mismas, sobre la relación entre la física clásica y la física cuántica, por favor, disculpen los reflejos condicionados de un profesor que comúnmente ha enseñado mecánica cuántica a estudiantes que no habían oído hablar de las matrices de espín de Pauli, hasta que llegaron a cursos avanzados. Y siempre parecía como si esperasen que la mecánica cuántica fuese una consecuencia lógica de la mecánica clásica que habían aprendido completamente en años anteriores. (Tal vez ellos querían evitar tener que aprender algo nuevo.) Ustedes han aprendido la fórmula clásica, ecuación (11.14), hace sólo unos pocos meses —y fue con advertencias de que era inadecuada— de modo que quizá ustedes no serán tan reticentes para aceptar la fórmula cuántica, ecuación (11.13), como verdad básica.

11-2 Las matrices de espín como operadores

Ya que estamos tratando la notación matemática, queremos describir *otra* manera para escribir las cosas —una manera que se usa con mucha frecuencia, porque es muy compacta—. Viene directamente de la notación introducida en el capítulo 8. Si tenemos un sistema en un estado $|\psi(t)\rangle$, que varía en el tiempo, podemos —como hicimos en la ecuación (8.34)— escribir la amplitud para que el sistema se encuentre en el estado $|i\rangle$ en $t + \Delta t$ en la forma

$$\langle i | \psi(t + \Delta t) \rangle = \sum_j \langle i | U(t + \Delta t, t) | j \rangle \langle j | \psi(t) \rangle$$

El elemento matricial $\langle i | U(t + \Delta t) | j \rangle$ es la amplitud para que el estado base $|j\rangle$ se convierta en el estado base $|i\rangle$ en el intervalo de tiempo Δt . Entonces *definimos* a H_{ij} escribiendo

$$\langle i | U(t + \Delta t, t) | j \rangle = \delta_{ij} - \frac{i}{\hbar} H_{ij}(t) \Delta t,$$

y demostramos que las amplitudes $C_i(t) = \langle i | \psi(t) \rangle$ se relacionan por las ecuaciones diferenciales

$$i\hbar \frac{dC_i}{dt} = \sum_j H_{ij} C_j. \quad (11.15)$$

Si escribimos las amplitudes C_i explícitamente, la misma ecuación aparece como

$$i\hbar \frac{d}{dt} \langle i | \psi \rangle = \sum_j H_{ij} \langle j | \psi \rangle. \quad (11.16)$$

Ahora, los elementos de matriz H_{ij} también son amplitudes que podemos escribir como $\langle i | H | j \rangle$; nuestra ecuación diferencial queda como

$$i\hbar \frac{d}{dt} \langle i | \psi \rangle = \sum_j \langle i | H | j \rangle \langle j | \psi \rangle. \quad (11.17)$$

Vemos que $-i/\hbar \langle i | H | j \rangle dt$ es la amplitud para que —bajo las condiciones físicas descritas por H — un estado $|j\rangle$, durante el intervalo de tiempo dt , “genere” el estado $|i\rangle$. (Todo esto está implícito en la discusión de la sección 8-4.)

Ahora, siguiendo las ideas de la sección 8-2, podemos eliminar el término común $\langle i |$ en la ecuación (11.7) —ya que es válida para cualquier estado $|i\rangle$ — y escribir la ecuación simplemente como

$$i\hbar \frac{d}{dt} |\psi\rangle = \sum_j H |j\rangle \langle j | \psi \rangle. \quad (11.18)$$

O, aún más, también podemos quitar j y escribir

$$i\hbar \frac{d}{dt} |\psi\rangle = H |\psi\rangle. \quad (11.19)$$

En el capítulo 8 señalamos que cuando escribíamos las cosas de este modo, la H en $H|j\rangle$ o en $H|\psi\rangle$ se llama *operador*. De ahora en adelante pondremos el sombrerito (^) sobre un operador para recordar que *es* un operador y no simplemente un número. Escribiremos $\hat{H}|\psi\rangle$. Aunque las dos ecuaciones (11.18) y (11.19) *significan exactamente lo mismo* que las ecuaciones (11.17) u (11.15), podemos *pensar* en ellas en una forma diferente. Por ejemplo, describiríamos a la ecuación (11.18) así: “La derivada respecto al tiempo del *vector de estado* $|\psi\rangle$ por $i\hbar$ es igual a lo que se obtiene aplicando el *operador* hamiltoniano $\hat{H}$ a cada estado base, multiplicado por la amplitud $\langle j | \psi \rangle$ para que ψ esté en el estado j , y sumando sobre todas las j ”. O la ecuación (11.19) se describe así: “La derivada respecto al tiempo (por $i\hbar$) de un estado $|\psi\rangle$ es igual a lo que se obtiene aplicando el hamiltoniano $\hat{H}$ al vector de estado $|\psi\rangle$ ”. Es sólo un modo abreviado de decir lo que está en la ecuación (11.17), pero, como verán, puede ser de gran conveniencia.

Si lo deseamos, podemos avanzar un paso más en la idea de “abstracción”. La ecuación (11.19) es válida para *cualquier estado* $|\psi\rangle$. También el lado izquierdo, $i\hbar d/dt$, es un operador —es la operación de “derivar respecto a t y multiplicar por $i\hbar$ ”. Por lo tanto, la ecuación (11.19) se puede considerar como una ecuación entre operadores: la ecuación operacional

$$i\hbar \frac{d}{dt} = \hat{H}.$$

El operador hamiltoniano (más o menos una constante) produce el mismo resultado que d/dt cuando actúa sobre cualquier estado. Recuerden que esta ecuación —como la ecuación (11.19)— *no* afirma que el operador $\hat{H}$ sea la *operación* idéntica a $i\hbar d/dt$. Las ecuaciones son la ley dinámica de la naturaleza —la ley del movimiento— para un sistema cuántico.

Sólo para obtener cierta práctica con estas ideas, les mostraremos otra forma para llegar a la ecuación (11.18). Saben que podemos escribir cualquier estado $|\psi\rangle$ en términos de sus proyecciones sobre algún conjunto de estados base [vean la ecuación (8.8)],

$$|\psi\rangle = \sum_i |i\rangle \langle i | \psi \rangle. \quad (11.20)$$

¿Cómo cambia $|\psi\rangle$ en el tiempo? Bien, simplemente derivando:

$$\frac{d}{dt} |\psi\rangle = \frac{d}{dt} \sum_i |i\rangle \langle i | \psi \rangle. \quad (11.21)$$

Ahora bien, los estados base $|i\rangle$ no varían en el tiempo (por lo menos *nosotros* siempre los consideramos como estados fijos definidos), pero las amplitudes $\langle i|\psi\rangle$ son números que pueden variar. Así, la ecuación (11.21) se convierte en

$$\frac{d}{dt}|\psi\rangle = \sum_i |i\rangle \frac{d}{dt}\langle i|\psi\rangle. \quad (11.22)$$

Ya que conocemos $d\langle i|\psi\rangle/dt$ de la ecuación (11.16), obtenemos

$$\begin{aligned} \frac{d}{dt}|\psi\rangle &= -\frac{i}{\hbar} \sum_i |i\rangle \sum_j H_{ij} \langle j|\psi\rangle \\ &= -\frac{i}{\hbar} \sum_{ij} |i\rangle \langle i|H|j\rangle \langle j|\psi\rangle = -\frac{i}{\hbar} \sum_j H|j\rangle \langle j|\psi\rangle. \end{aligned}$$

Ésta es de nuevo la ecuación (11.18).

Así entonces, tenemos muchas maneras para considerar al hamiltoniano. Podemos pensar que el conjunto de coeficientes H_{ij} es simplemente un montón de números, o podemos pensar en las “amplitudes” $\langle i|H|j\rangle$, o podemos pensar en la “matriz” H_{ij} , o podemos pensar en el “operador” $\hat{H}$. Todo significa lo mismo.

Ahora volvamos a nuestros sistemas de dos estados. Si escribimos el hamiltoniano en términos de las matrices sigma (con coeficientes numéricos apropiados, como B_x , etc.), claramente podemos considerar también a σ_{ij}^x como una amplitud $\langle i|\sigma_x|j\rangle$ o, para abreviar, como el operador $\hat{\sigma}_x$. Si usamos los operadores, entonces podemos escribir la ecuación de movimiento de un estado $|\psi\rangle$ en un campo magnético como

$$i\hbar \frac{d}{dt}|\psi\rangle = -\mu(B_x\hat{\sigma}_x + B_y\hat{\sigma}_y + B_z\hat{\sigma}_z)|\psi\rangle. \quad (11.23)$$

Cuando queramos “usar” esa ecuación, tendremos que expresar a $|\psi\rangle$ en términos de los vectores base (del mismo modo que tenemos que encontrar las componentes de los vectores espaciales cuando queremos números específicos). Así que, usualmente, queremos escribir la ecuación (11.23) en la forma un tanto desarrollada:

$$i\hbar \frac{d}{dt}|\psi\rangle = -\mu \sum_i (B_x\hat{\sigma}_x + B_y\hat{\sigma}_y + B_z\hat{\sigma}_z)|i\rangle \langle i|\psi\rangle. \quad (11.24)$$

Ahora verán por qué la idea de operador es tan ingeniosa. Para usar la ecuación (11.24), necesitamos saber lo que sucede cuando los operadores $\hat{\sigma}$ se aplican a cada uno de los estados base. Averigüémoslo. Supongan que tenemos $\hat{\sigma}_z|+\rangle$; es algún vector $|?\rangle$, pero ¿cuál? Bien, multipliquémoslo por la izquierda por $\langle +|$; tenemos:

$$\langle +|\hat{\sigma}_z|+\rangle = \sigma_{11}^z = 1$$

(usando la tabla 11-1). Así sabemos que

$$\langle +|?\rangle = 1. \quad (11.25)$$

Ahora multipliquemos a $\hat{\sigma}_z|+\rangle$ por la izquierda por $\langle -|$. Obtenemos

$$\langle -|\hat{\sigma}_z|+\rangle = \sigma_{21}^z = 0;$$

y por lo tanto

$$\langle -|?\rangle = 0. \quad (11.26)$$

Hay solamente un vector de estado que satisface a la vez (11.25) y (11.26); es $|+\rangle$. Descubrimos entonces que

$$\hat{\sigma}_z|+\rangle = |+\rangle. \quad (11.27)$$

Tabla 11-3

Propiedades del operador $\hat{\sigma}$

$\hat{\sigma}_z +\rangle = +\rangle$
$\hat{\sigma}_z -\rangle = - -\rangle$
$\hat{\sigma}_x +\rangle = -\rangle$
$\hat{\sigma}_x -\rangle = +\rangle$
$\hat{\sigma}_y +\rangle = i -\rangle$
$\hat{\sigma}_y -\rangle = -i +\rangle$

Con esta clase de razonamiento pueden demostrar fácilmente que todas las propiedades de las matrices sigma se pueden describir en la notación operacional mediante el conjunto de reglas dadas en la tabla 11-3.

Si tenemos productos de matrices sigma, se convierten en productos de operadores. Cuando dos operadores aparecen juntos como un producto, primero efectúen la operación con el operador que se encuentra más a la derecha. Por ejemplo, por $\hat{\sigma}_x \hat{\sigma}_y |+\rangle$ tenemos que entender $\hat{\sigma}_x(\hat{\sigma}_y |+\rangle)$. De la tabla 11-3, tenemos que $\hat{\sigma}_y |+\rangle = i|-\rangle$, y así

$$\hat{\sigma}_x \hat{\sigma}_y |+\rangle = \hat{\sigma}_x(i|-\rangle). \quad (11.28)$$

Pero cualquier número —como i — pasa al otro lado de un operador (los operadores sólo funcionan sobre vectores de estado); por lo que la ecuación (11.28) es lo mismo que

$$\hat{\sigma}_x \hat{\sigma}_y |+\rangle = i \hat{\sigma}_x |-\rangle = i |+\rangle.$$

Si hacen lo mismo para $\hat{\sigma}_x \hat{\sigma}_y |-\rangle$, encontrarán que

$$\hat{\sigma}_x \hat{\sigma}_y |-\rangle = -i |-\rangle.$$

Al observar la tabla 11-3, ven que aplicar $\hat{\sigma}_x \hat{\sigma}_y$ a $|+\rangle$ o a $|-\rangle$ da exactamente lo mismo que obtendrían si aplican a $\hat{\sigma}_z$ y multiplican por i . Por lo tanto, podemos decir que la operación $\hat{\sigma}_x \hat{\sigma}_y$ es idéntica a la operación $i\hat{\sigma}_z$, y escribir este enunciado en forma de ecuación operacional:

$$\hat{\sigma}_x \hat{\sigma}_y = i\hat{\sigma}_z. \quad (11.29)$$

Observen que esta ecuación es idéntica a una de las ecuaciones matriciales de la tabla 11-2. Así, encontramos de nuevo la correspondencia entre los puntos de vista matricial y operacional. Por lo tanto, cada una de las ecuaciones de la tabla 11-2 también se puede considerar como una ecuación de operadores sigma. Pueden comprobar que verdaderamente resultan de la tabla 11-3. Es mejor, cuando trabajamos con estas cosas, *no* insistir en si una cantidad como σ o H es una matriz o un operador. Todas las ecuaciones son iguales en una forma o en la otra, por eso, la tabla 11-2 sirve para los operadores sigma, o para las matrices sigma, según deseen.

11-3 La solución de las ecuaciones de dos estados

Ahora podemos escribir nuestra ecuación para dos estados en varias formas, por ejemplo, ya sea como

$$i\hbar \frac{dC_i}{dt} = \sum_j H_{ij} C_j$$

o bien como

$$i\hbar \frac{d|\psi\rangle}{dt} = \hat{H} |\psi\rangle. \quad (11.30)$$

Ambas significan lo mismo. Para una partícula de espín un medio en un campo magnético, el hamiltoniano H está dado por la ecuación (11.8) o por la ecuación (11.13).

Si el campo está en la dirección z , entonces —como ya hemos visto varias veces— la solución es que el estado $|\psi\rangle$, cualquiera que sea, precesa alrededor del eje z (como si tomasen al objeto físico y lo girasen alrededor del eje z) con una velocidad angular igual a dos veces el campo magnético por $\mu/\hbar$. Lo mismo es verdad, naturalmente, para un campo magnético en cualquier otra dirección, ya que la física es independiente del sistema de coordenadas. Si tenemos una situación donde el campo magnético varía de un instante a otro de un modo muy complicado, podemos analizar esta situación de la siguiente forma. Supongan que empiezan con el espín en la dirección $+z$ y tienen un campo magnético en x . El espín comienza a girar. Si el campo en x se apaga, el espín deja de girar. Ahora, si encendemos un campo en la dirección z , el espín precesa alrededor

de z , y así podríamos seguir. Por consiguiente, dependiendo de cómo el campo varía en el tiempo, pueden calcular cuál es el estado final —en la dirección de qué eje apuntará—. Entonces pueden referir ese estado a los originales $|+\rangle$ y $|-\rangle$ respecto a z , usando las fórmulas de proyección que teníamos en el capítulo 10 (o en el capítulo 6). Si el estado termina con su espín en la dirección (θ, ϕ) , tendría una amplitud $\cos(\theta/2)e^{-i\phi/2}$ para estar hacia arriba y una amplitud $\sin(\theta/2)e^{+i\phi/2}$ para estar hacia abajo. Esto resuelve cualquier problema. Es una descripción verbal de la solución de las ecuaciones diferenciales.

La solución que acabamos de describir es suficientemente general para tratar *cualquier sistema de dos estados*. Tomemos nuestro ejemplo de la molécula de amoníaco —incluyendo los efectos de un campo eléctrico—. Si describimos el sistema en términos de los estados $|I\rangle$ y $|II\rangle$, las ecuaciones (9.38) y (9.39) se parecerán a:

$$\begin{aligned} i\hbar \frac{dC_I}{dt} &= +AC_I + \mu\mathcal{E}C_{II}, \\ i\hbar \frac{dC_{II}}{dt} &= -AC_{II} + \mu\mathcal{E}C_I. \end{aligned} \quad (11.31)$$

Ustedes dirán, “No, me acuerdo que había un E_0 en ellas”. Bien, hemos cambiado el origen de energía de modo que E_0 sea cero. (Siempre pueden hacer esto, cambiando las dos amplitudes en el mismo factor $-e^{iE_0T/\hbar}$ y se libran de cualquier energía constante.) Ahora bien, si las ecuaciones correspondientes tienen siempre las mismas soluciones, no tenemos realmente que hacerlo dos veces. Si observamos estas ecuaciones y a las ecuaciones (11.1), entonces podemos hacer la siguiente identificación. Llamemos $|I\rangle$ al estado $|+\rangle$ y $|II\rangle$ al estado $|-\rangle$. Esto *no significa* que estemos alineando al amoníaco en el espacio, o que $|+\rangle$ y $|-\rangle$ tengan algo que ver con el eje z . Es meramente artificial. Tenemos un espacio artificial que podríamos llamar el “espacio representativo de la molécula de amoníaco”, o cualquier otra cosa —un “diagrama” tridimensional en el cual estar “hacia arriba” corresponde a tener la molécula en el estado $|I\rangle$, y estar “hacia abajo” en la dirección de este falso eje z representa tener a la molécula en el estado $|II\rangle$ —. Entonces, las ecuaciones se identificarán como sigue. Primero que nada, ven que el hamiltoniano se puede escribir en términos de las matrices sigma en la forma

$$H = +A\sigma_z + \mu\mathcal{E}\sigma_x. \quad (11.32)$$

Dicho de otro modo, μB_z en la ecuación (11.1) corresponde a $-A$ en la ecuación (11.32), y μB_x corresponde a $-\mu\mathcal{E}$. En nuestro espacio “modelo”, entonces, tenemos un campo constante B en la dirección z . Si tenemos un campo eléctrico $\mathcal{E}$ que varía con el tiempo, entonces tendremos un campo B en la dirección x que varía proporcionalmente. *Por lo que el comportamiento de un electrón en un campo magnético que tiene una componente constante en la dirección z y una componente oscilante en la dirección x , es matemáticamente análogo, y corresponde exactamente al comportamiento de una molécula de amoníaco en un campo eléctrico oscilante.* Desafortunadamente, no tenemos tiempo para entrar en los detalles de esta correspondencia, o para calcular cualquiera de los detalles técnicos. Solamente deseábamos señalar que *todos* los sistemas de dos estados se pueden tratar en forma análoga a un objeto de espín un medio que precesas en un campo magnético.

11-4 Los estados de polarización del fotón

Hay otros sistemas de dos estados que son interesantes para estudiar, y el primero del que queremos hablar es el fotón. Para describir a un fotón, primero debemos dar su vector de momentum. Para un fotón libre, la frecuencia está determinada por el momentum, por lo tanto no tenemos que decir además cuál es su frecuencia. No obstante, existe también una propiedad llamada polarización. Imaginen que un fotón se acerca a ustedes con una frecuencia monocromática definida (la que conservaremos a través de toda la discusión, y así no tendremos una variedad de estados de momentum). Entonces hay dos

direcciones de polarización. En la teoría clásica, la luz se puede describir como que posee un campo eléctrico que oscila horizontal o verticalmente (por ejemplo); estas dos clases de luz se llaman luz polarizada en la dirección x y luz polarizada en la dirección y . La luz también puede estar polarizada en cualquier otra dirección, que se puede construir por la superposición del campo eléctrico en la dirección x y el otro en la y . O si toman las componentes x y y fuera de fase en 90° , obtienen un campo eléctrico que rota —la luz está elípticamente polarizada—. (Esto es un recuento rápido de la teoría clásica de la luz polarizada que estudiamos en el capítulo 33 del volumen I.)

Supongan ahora, sin embargo, que tenemos un *solo* fotón —y solamente uno—. No hay campo eléctrico que podamos discutir en la misma forma. Todo lo que tenemos es *un fotón*. Pero un fotón tiene que tener lo análogo a los fenómenos clásicos de polarización. Debe haber por lo menos dos clases diferentes de fotones. En principio, podrían pensar que debería haber una variedad infinita —después de todo, el vector eléctrico puede apuntar en toda clase de direcciones—. Sin embargo, podemos describir la polarización de un fotón como un sistema de dos estados. Un fotón puede estar en el estado $|x\rangle$ o en el $|y\rangle$. Por $|x\rangle$ entendemos el estado de polarización de cada uno de los fotones en un haz de luz que está polarizado de manera *clásica* en la dirección x . Por otra parte, por $|y\rangle$ entendemos el estado de polarización de cada uno de los fotones de un haz polarizado en la dirección y . Y podemos tomar a $|x\rangle$ y $|y\rangle$ como los estados base de un fotón de momentum dado que apunta hacia ustedes —que llamaremos dirección z —. Por lo que hay dos estados base $|x\rangle$ y $|y\rangle$, y esto es todo lo que necesitamos para describir cualquier fotón.

Por ejemplo, si tenemos un polarizador dispuesto con su eje para permitir pasar luz polarizada en la dirección que llamamos x , y enviamos un fotón que está en el estado $|y\rangle$, éste será absorbido por el polarizador. Si enviamos un fotón que está en el estado $|x\rangle$, pasará y seguirá siendo $|x\rangle$. Si tomamos una lámina de calcita que recibe un haz de luz polarizada y los divide en un haz $|x\rangle$ y un haz $|y\rangle$, esa lámina de calcita será la perfecta analogía del aparato de Stern-Gerlach, que separa un haz de átomos de plata en los estados $|+\rangle$ y $|-\rangle$. Entonces, todo lo que hicimos antes con partículas y aparatos de Stern-Gerlach, podemos hacerlo de nuevo con luz y láminas de calcita. ¿Y qué ocurre con la luz filtrada a través de un polarizador dispuesto en un ángulo θ ? ¿Es ese otro estado? Sí, en realidad, *es* otro estado. Llamemos x' al eje del polarizador para distinguirlo de los ejes de nuestros estados base. Vean la figura 11-2. Cualquier fotón que salga estará en el estado $|x'\rangle$. Sin embargo, cualquier estado se puede representar como una combinación lineal de los estados base, y la fórmula de la combinación es, en este caso,

$$|x'\rangle = \cos \theta |x\rangle + \sin \theta |y\rangle. \quad (11.33)$$

Esto es, si un fotón atraviesa un polarizador puesto en un ángulo θ (respecto a x), aun así se puede separar en haces $|x\rangle$ y $|y\rangle$ —mediante una lámina de calcita, por ejemplo—. O pueden, si lo desean, sólo descomponerlo en las componentes x y y en su imaginación. De las dos formas, siempre encontrarán las amplitudes $\cos \theta$ para estar en el estado $|x\rangle$ y la amplitud $\sin \theta$ para estar en el estado $|y\rangle$.

Ahora nos preguntamos: supongan que se polariza un fotón en la dirección x' mediante un polarizador dispuesto en un ángulo θ , y llega a un polarizador en un ángulo cero —como en la figura 11-3—; ¿qué sucederá? ¿Con qué probabilidad atravesará? La respuesta es la siguiente. Después que atraviesa el primer polarizador, definitivamente se encuentra en el estado $|x'\rangle$. El segundo polarizador dejará pasar al fotón, si éste se encuentra en el estado $|x\rangle$ (pero lo absorberá si está en el estado $|y\rangle$). Por eso estamos preguntando ¿con qué probabilidad el fotón aparece en el estado $|x\rangle$? Obtenemos esta probabilidad del cuadrado de la norma de la amplitud $\langle x | x' \rangle$ para que un fotón en el estado $|x'\rangle$ esté también en el estado $|x\rangle$. ¿Qué es $\langle x | x' \rangle$? Sólo multipliquen a la ecuación (11.33) por $\langle x |$ para obtener

$$\langle x | x' \rangle = \cos \theta \langle x | x \rangle + \sin \theta \langle x | y \rangle.$$

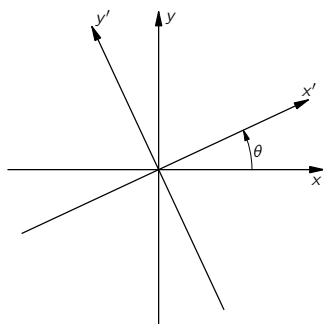
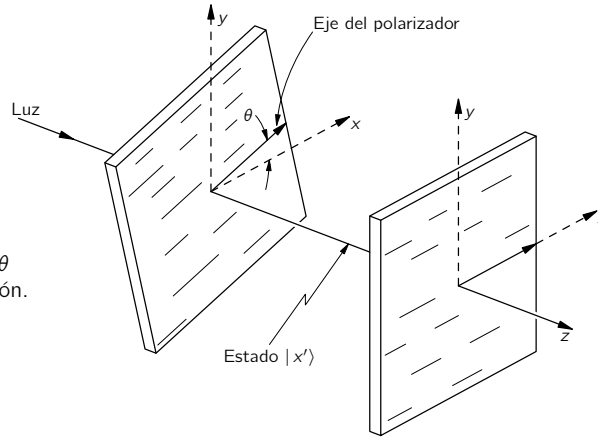


Figura 11-2. Coordenadas perpendiculares al vector de momentum del fotón.

Figura 11-3. Dos láminas polarizadoras con un ángulo θ entre los planos de polarización.



Ahora bien, $\langle x | y \rangle = 0$, por cuestiones físicas —como *debe* ser si $|x\rangle$ y $|y\rangle$ son estados base— y $\langle x | x \rangle = 1$. Nos queda entonces

$$\langle x | x' \rangle = \cos \theta,$$

y la probabilidad es $\cos^2 \theta$. Por ejemplo, si el primer polarizador está colocado a 30° , un fotón atravesará los $3/4$ de las veces, y $1/4$ de las veces calentará el polarizador al ser absorbido en él.

Ahora veamos lo que sucede clásicamente en la misma situación. Tenemos un haz de luz con un campo eléctrico que varía de un modo u otro —digamos “sin polarizar”—. Después de atravesar el primer polarizador, el campo eléctrico estará oscilando en la dirección x' , con un tamaño $\mathcal{E}$; representamos al campo como un vector oscilante con un valor máximo $\mathcal{E}_0$ como en la figura 11-4. Ahora bien, cuando la luz llega al segundo polarizador, sólo la componente x , $\mathcal{E}_0 \cos \theta$, del campo eléctrico logra pasar. La *intensidad* es proporcional al cuadrado del campo y, por lo tanto, a $\mathcal{E}_0^2 \cos^2 \theta$. Y así la energía que atraviesa es $\cos^2 \theta$ veces más débil que la energía que entra en el último polarizador.

Las imágenes clásica y cuántica dan resultados semejantes. Si lanzasen 10 000 millones de fotones al segundo polarizador, la probabilidad promedio para que cada uno atravesase es, digamos, $3/4$, esperarían que $3/4$ de los 10 000 millones pasarán. Del mismo modo, la energía que transportan sería $3/4$ de la energía que ustedes trataban de hacer pasar. La teoría clásica no dice nada sobre la estadística de esto —dice sencillamente que la energía que atraviesa es precisamente $3/4$ de la energía que enviaron—. Esto es, por supuesto, imposible si solamente hay un fotón. No hay tal cosa como $3/4$ de un fotón. Ya sea que está *por completo*, o no lo está. La mecánica cuántica nos dice que está *por completo* ahí $3/4$ de las veces. La relación entre las dos teorías es clara.

¿Y qué pasa con las otras clases de polarización? Por ejemplo, ¿con la polarización circular derecha? En la teoría clásica, la polarización circular derecha tiene componentes iguales en x y en y , que están 90° fuera de fase. En la teoría cuántica, un fotón polarizado circularmente a la derecha (RHC, por sus siglas en inglés) tiene amplitudes iguales para estar polarizado en $|x\rangle$ o en $|y\rangle$, y las *amplitudes* están 90° fuera de fase. Si llamamos a un fotón RHC el estado $|R\rangle$, y a un fotón LHC el estado $|L\rangle$, podemos escribir (vean el volumen I, sección 33-1)

$$\begin{aligned} |R\rangle &= \frac{1}{\sqrt{2}} (|x\rangle + i|y\rangle), \\ |L\rangle &= \frac{1}{\sqrt{2}} (|x\rangle - i|y\rangle) \end{aligned} \tag{11.34}$$

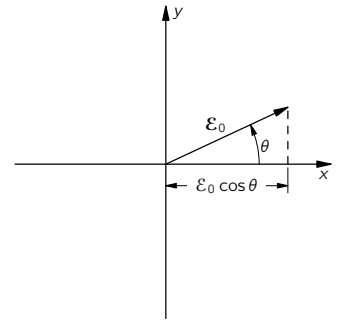


Figura 11-4. Representación clásica del vector eléctrico $\mathcal{E}$.

—donde se ha puesto el $1/\sqrt{2}$ para tener estados normalizados—. Con estos estados pueden calcular los efectos de filtración o de interferencia que deseen,

usando las leyes de la teoría cuántica. Si quieren, también pueden escoger a $|R\rangle$ y $|L\rangle$ como estados base, y representar todo en función de ellos. Solamente necesitan demostrar primero que $\langle R|L\rangle = 0$ —lo que pueden hacer tomando la forma conjugada de la primera ecuación anterior [vean la ecuación (8.13)] y multiplicándola por la otra—. Pueden descomponer la luz usando polarizaciones en x y y , o en polarizaciones x' y y' , o polarizaciones derecha e izquierda como base.

Sólo como ejemplo, tratemos de dar vuelta a nuestras fórmulas. ¿Podemos representar al estado $|x\rangle$ como una combinación lineal de derecha e izquierda? Sí, aquí está:

$$\begin{aligned}|x\rangle &= \frac{1}{\sqrt{2}}(|R\rangle + |L\rangle), \\ |y\rangle &= -\frac{i}{\sqrt{2}}(|R\rangle - |L\rangle).\end{aligned}\tag{11.35}$$

Demostración: Sumen y resten las dos ecuaciones en (11.34). Es fácil pasar de una base a otra.

No obstante, tenemos que señalar un punto curioso. Si un fotón está polarizado circularmente a la derecha, no debería tener nada que ver con los ejes x y y . Si considerásemos lo mismo desde un sistema de coordenadas girado un cierto ángulo alrededor de la dirección de propagación, la luz aún estaría polarizada circularmente a la derecha —y de manera similar para la izquierda—. La luz polarizada circularmente a la derecha o la izquierda es la misma para cualquier rotación de ese tipo; la definición es independiente de la dirección x que escojamos (excepto que la dirección del fotón está dada). ¿No es interesante? —no necesita de ejes para su definición—. Mucho mejor que x y y . Por otra parte, ¿no es casi un milagro que *sumando* las polarizaciones derecha e izquierda puedan encontrar cuál era la dirección x ? Si “derecha” e “izquierda” no dependen en ningún modo de x , ¿por qué es que podemos colocándolas juntas obtener x ? Podemos, en parte, responder esta pregunta especificando el estado $|R'\rangle$, que representa a un fotón con polarización RHC en el sistema x' , y' . En este sistema, escribirían

$$|R'\rangle = \frac{1}{\sqrt{2}}(|x'\rangle + i|y'\rangle).$$

¿Cómo se ve ese estado en el sistema x , y ? Sustituyan simplemente $|x'\rangle$ de la ecuación (11.33) y el $|y'\rangle$ correspondiente —no lo escribimos antes, pero es $(-\sin\theta)|x\rangle + (\cos\theta)|y\rangle$ —. Luego entonces,

$$\begin{aligned}|R'\rangle &= \frac{1}{\sqrt{2}}[\cos\theta|x\rangle + \sin\theta|y\rangle - i\sin\theta|x\rangle + i\cos\theta|y\rangle] \\ &= \frac{1}{\sqrt{2}}[(\cos\theta - i\sin\theta)|x\rangle + i(\cos\theta - i\sin\theta)|y\rangle] \\ &= \frac{1}{\sqrt{2}}(|x\rangle + i|y\rangle)(\cos\theta - i\sin\theta).\end{aligned}$$

El primer término es justamente $|R\rangle$, y el segundo es $e^{-i\theta}$; nuestro resultado es que

$$|R'\rangle = e^{-i\theta}|R\rangle.\tag{11.36}$$

Los estados $|R'\rangle$ y $|R\rangle$ son iguales, excepto por el factor de fase $e^{-i\theta}$. Si hacen lo mismo para $|L'\rangle$, obtienen†

$$|L'\rangle = e^{+i\theta}|L\rangle.\tag{11.37}$$

† Es muy similar a lo que encontramos (en el capítulo 6) para una partícula con espín un medio cuando girábamos las coordenadas alrededor del eje z —entonces obtuvimos los factores de fase $e^{\pm i\phi/2}$ —. Es, de hecho, exactamente lo que escribimos en la sección 5-7 para los estados $|+\rangle$ y $|-\rangle$ de una partícula con espín uno —lo cual no es coincidencia—. El fotón es una partícula con espín uno que, sin embargo, no tiene estado “cero”.

Ahora ven lo que sucede. Si sumamos $|R\rangle$ y $|L\rangle$, obtenemos algo diferente de lo que obtenemos al sumar $|R'\rangle$ y $|L'\rangle$. Por ejemplo, un fotón polarizado en x es [ecuación (11.35)] la suma de $|R\rangle$ y $|L\rangle$, y uno polarizado en y es la suma pero con la fase de uno adelantada 90° y la del otro atrasada 90° . Esto es exactamente lo que obtenemos de la suma de $|R'\rangle$ y $|L'\rangle$ para el ángulo especial $\theta = 90^\circ$, y ya está. Una polarización en x para el sistema *prima* es igual a una polarización en y para el sistema original. Por esto, no es exactamente cierto que un fotón polarizado circularmente se vea igual en cualquier conjunto de ejes. Su *fase* (la relación de fase entre los estados polarizados circularmente a la derecha e izquierda) conserva la información de la dirección x .

11-5 El mesón K neutro†

Ahora describiremos un sistema de dos estados en el mundo de las partículas extrañas —un sistema para el cual la mecánica cuántica da una predicción muy notable—. Para describirlo por completo nos involucraríamos con muchas cosas acerca de las partículas extrañas, por lo que, desafortunadamente, tendremos que recortar bastante. Sólo podemos dar una idea general sobre cómo se llevó a cabo cierto descubrimiento —para mostrarles la clase de razonamiento que estaba involucrado—. Comienza con el descubrimiento por Gell-Mann y Nishijima del concepto de *extrañeza* y de una nueva ley de *conservación de la extrañeza*. Fue cuando Gell-Mann y Pais analizaban las consecuencias de estas nuevas ideas, que tropezaron con la predicción de un fenómeno muy notable que vamos a describir. En primer lugar, sin embargo, tenemos que hablar un poco acerca de la “extrañeza”.

Debemos comenzar con lo que se llama *interacciones fuertes* de las partículas nucleares. Éstas son las interacciones responsables de las intensas fuerzas nucleares —tan distintas, por ejemplo, de las interacciones electromagnéticas relativamente más débiles—. Las interacciones son “fuertes” en el sentido de que si dos partículas se acercan lo suficiente como para interactuar, entonces interactúan de modo notable, y producen otras partículas muy fácilmente. Las partículas nucleares también tienen lo que se llama una “interacción débil”, por medio de la cual pueden suceder ciertas cosas, como la desintegración beta, pero siempre muy despacio en una escala temporal nuclear —las interacciones débiles son muchos, pero muchos órdenes de magnitud más débiles que las interacciones fuertes, y aún más débiles que las electromagnéticas—.

Cuando se estudiaban las interacciones fuertes con los grandes aceleradores, muchos se sorprendieron al descubrir que ciertas cosas que “deberían” ocurrir —que se esperaba que sucedieran— no ocurrieron. Por ejemplo, en algunas interacciones, una partícula de un cierto tipo no aparecía cuando era esperada. Gell-Mann y Nishijima notaron que muchos de estos sucesos peculiares se podrían explicar de una vez si se inventaba una nueva ley de conservación: la *conservación de la extrañeza*. Propusieron que había una nueva clase de atributo asociado con cada partícula —que llamaron su número de “extrañeza”— y que en cualquier interacción fuerte la “cantidad de extrañeza” se conserva.

Supongan, por ejemplo, que un mesón K negativo de alta energía —con una energía de GeV— choca con un protón. De la interacción pueden surgir muchas otras partículas: mesones π , mesones K, partículas lambda, partículas sigma —cualquiera de los mesones o bariones listados en la tabla 2-2 del volumen I—. Sin embargo, se observó que sólo aparecen *ciertas combinaciones*, y nunca las otras. Ahora bien, se conocía que ciertas leyes de conservación se debían cumplir. En primer lugar, la energía y el momentum siempre se conservan. La energía total y el momentum después de un suceso deben ser iguales que antes. En

† Nos damos cuenta ahora que el material de esta sección es más largo y más difícil de lo que es conveniente a este nivel en nuestro desarrollo. Sugerimos que lo omitan y continúen con la sección 11-6. Si son ambiciosos y tienen tiempo, quizá deseen volver a éste más adelante. Lo dejamos aquí, porque es un lindo ejemplo —tomado de un trabajo reciente en física de altas energías— de lo que se puede hacer con nuestra formulación cuántica de los sistemas de dos estados.

segundo lugar, se tiene la conservación de la carga eléctrica, que dice que la carga total de las partículas que salen debe ser igual a la carga total que transportan las partículas originales. En nuestro ejemplo de la colisión de un mesón K y de un protón, las siguientes reacciones *sí* ocurren:

$$\begin{aligned} & \text{K}^- + \text{p} \rightarrow \text{p} + \text{K}^- + \pi^+ + \pi^- + \pi^0 \\ \text{o} \end{aligned} \tag{11.38}$$

$$\text{K}^- + \text{p} \rightarrow \Sigma^- + \pi^+.$$

Nunca obtendríamos:

$$\text{K}^- + \text{p} \rightarrow \text{p} + \text{K}^- + \pi^+ \quad \text{o} \quad \text{K}^- + \text{p} \rightarrow \Lambda^0 + \pi^+, \tag{11.39}$$

debido a la conservación de la carga. También se sabía que se conserva el *número de bariones*. El número de bariones que *sale* debe ser igual al número de bariones que *entra*. En esta ley, una *antipartícula* de un barión se cuenta como *menos* un barión. Esto significa que podemos ver —y en realidad *sí* vemos—

$$\begin{aligned} & \text{K}^- + \text{p} \rightarrow \Lambda^0 + \pi^0 \\ \text{o} \end{aligned} \tag{11.40}$$

$$\text{K}^- + \text{p} \rightarrow \text{p} + \text{K}^- + \text{p} + \bar{\text{p}}$$

(donde $\bar{\text{p}}$ es el antiprotón, el cual tiene una carga negativa). Pero *nunca* vemos

$$\begin{aligned} & \text{K}^- + \text{p} \rightarrow \text{K}^- + \pi^+ + \pi^0 \\ \text{ni} \end{aligned} \tag{11.41}$$

$$\text{K}^- + \text{p} \rightarrow \text{p} + \text{K}^- + \text{n}$$

(aun cuando haya mucha energía), porque no se conservarían los bariones.

Estas leyes, sin embargo, *no* explican el hecho extraño de que las siguientes reacciones —las cuales no parece que sean, a primera vista, especialmente diferentes de algunas en (11.38) o (11.40)— tampoco se hayan observado:

$$\begin{aligned} & \text{K}^- + \text{p} \rightarrow \text{p} + \text{K}^- + \text{K}^0 \\ \text{o} \end{aligned}$$

$$\text{K}^- + \text{p} \rightarrow \text{p} + \pi^- \tag{11.42}$$

o

$$\text{K}^- + \text{p} \rightarrow \Lambda^0 + \text{K}^0.$$

La explicación es la conservación de la extrañeza. Cada partícula está acompañada de un número —su *extrañeza* S — y hay una ley que, para cualquier interacción *fuerte*, la extrañeza total que *sale* debe ser igual a la extrañeza total que *entró*. El protón y el antiprotón (p , $\bar{\text{p}}$), el neutrón y el antineutrón (n , $\bar{\text{n}}$) y los mesones π (π^+ , π^0 , π^-) tienen número de extrañeza *cero*; por otro lado, los mesones K^+ y K^0 tienen extrañeza $+1$, los K^- y $\bar{\text{K}}^0$ (anti- K^0),[†] las partículas Λ^0 y Σ ($+$, 0 , $-$) tienen extrañeza -1 . Hay también una partícula con extrañeza -2 —la partícula Ξ (“Xi” mayúscula)— y tal vez otras aún desconocidas. Hemos hecho una lista de estas partículas extrañas en la tabla 11-4.

Veamos cómo se aplica la conservación de la extrañeza en algunas de las reacciones que hemos escrito. Si empezamos con un K^- y un protón, tenemos una extrañeza total de $(-1 + 0) = -1$. La conservación de la extrañeza dice que la suma de la extrañeza de los productos *después* de la reacción también debe ser -1 . Ven que ocurre eso para las reacciones (11.38) y (11.40). Pero en las reacciones de (11.42), la extrañeza del lado derecho es *cero* en cada caso, las cuales no conservan la extrañeza, y no ocurren. ¿Por qué? Nadie lo sabe. Nadie sabe más de lo que nosotros hemos dicho. La naturaleza funciona de esta forma.

Ahora consideremos la siguiente reacción: un π^- choca con un protón. Podríamos obtener, por ejemplo, una partícula Λ^0 más una partícula K neutra

[†] Léase: “K-cero barra”.

Tabla 11-4

Números de extrañeza de las partículas en interacciones fuertes

	S			
	-2	-1	0	+1
Bariones	Ξ^0 Ξ^-	Σ^+ Λ^0, Σ^0 Σ^-	p n	
Mesones		$\bar{K}^0$ K^-	π^+ π^0 π^-	K^+ K^0

Nota: π^- es la antipartícula de π^+ (y *viceversa*).

—dos partículas neutras—. Ahora bien, ¿qué K neutra obtienen? Puesto que la partícula Λ tiene extrañeza -1 y la π y el p^+ tienen extrañeza cero, y puesto que ésta es una reacción de producción rápida, la extrañeza no debe cambiar. La partícula K debe tener extrañeza $+1$ —por lo tanto debe ser el K^0 —. La reacción es

$$\pi^- + p \rightarrow \Lambda^0 + K^0,$$

con

$$S = 0 + 0 = -1 + +1 \quad (\text{se conserva}).$$

Si se obtuviera $\bar{K}^0$ en lugar de K^0 , la extrañeza a la derecha sería -2 —que la naturaleza no permite, ya que la extrañeza del lado izquierdo es cero—. Por otra parte, un $\bar{K}^0$ se puede obtener en otras reacciones, tales como

$$n + n \rightarrow n + \bar{p} + \bar{K}^0 + K^+,$$

$$S = 0 + 0 = 0 + 0 + -1 + +1$$

o

$$K^- + p \rightarrow n + \bar{K}^0,$$

$$S = -1 + 0 = 0 + -1.$$

Puede que estén pensando, “esto es demasiado, porque ¿cómo *saben* si es un $\bar{K}^0$ o un K^0 ? Se ven exactamente igual. Son antipartículas una de la otra, por lo que tienen exactamente la misma masa, y ambas tienen carga eléctrica cero. ¿Cómo las distinguen?” Por las reacciones que *éstas* producen. Por ejemplo, un $\bar{K}^0$ puede interactuar con materia y producir una partícula Λ , de este modo:

$$\bar{K}^0 + p \rightarrow \Lambda^0 + \pi^+,$$

pero un K^0 no puede. *No* hay manera de que un K^0 pueda producir una partícula Λ cuando interactúa con materia ordinaria (protones y neutrones).[†] Por consiguiente, la distinción experimental entre K^0 y $\bar{K}^0$ sería que una de ellas producirá partículas Λ y la otra no.

Una de las predicciones de la teoría de la extrañeza es ésta —si, en un experimento con piones de alta energía, se produce una partícula Λ con un mesón K neutro, entonces cuando *ese* mesón K neutro llegue a otros pedazos de materia nunca producirá de nuevo una Λ —. El experimento se podría hacer en una forma algo parecida a ésta. Envían un haz de mesones π^- dentro de una gran cámara con burbujas de hidrógeno. El rastro del π^- desaparece, pero en alguna otra parte aparecen un par de rastros (un protón y un π^-) indicando que una

[†] Excepto, naturalmente, si *también* produce *dos* K^+ u otras partículas con una extrañeza total de $+2$. Aquí podemos pensar en reacciones en las que no hay suficiente energía para producir estas partículas extrañas adicionales.

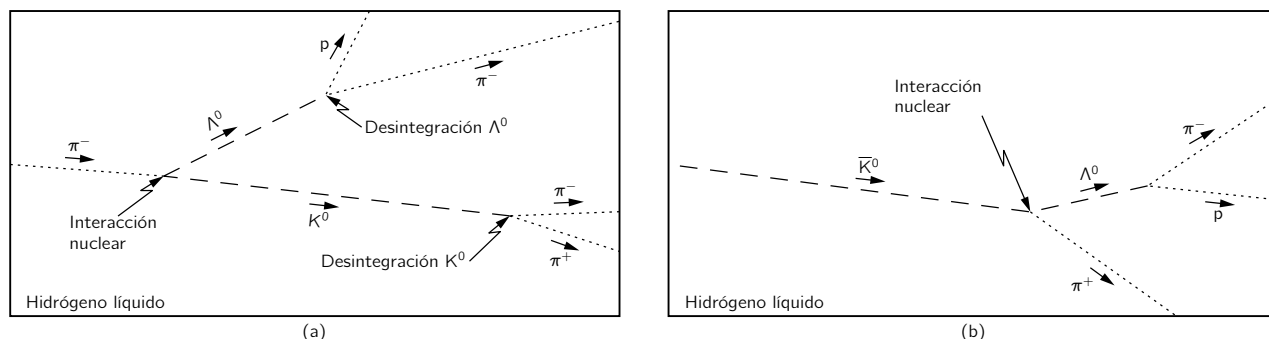


Figura 11-5. Eventos de alta energía tal como se observan en una cámara de burbujas de hidrógeno. (a) Un mesón π^- interactúa con un núcleo de hidrógeno (protón) produciendo una partícula Λ^0 y un mesón K^0 . Ambas partículas se desintegran en la cámara. (b) Un mesón $\bar{K}^0$ interactúa con un protón produciendo un mesón π^- y una partícula Λ^0 que luego se desintegra. (Las partículas neutras no dejan rastro. Sus trayectorias supuestas están indicadas aquí mediante líneas punteadas.)

partícula Λ se ha desintegrado† —véase la figura 11-5—. Entonces saben que en alguna parte hay una K^0 que no pueden ver.

Pueden, no obstante, determinar por dónde va usando la conservación del momentum y de la energía. [Se podría revelar más tarde, al desintegrarse en dos partículas cargadas, como muestra la figura 11-5(a).] Cuando el K^0 está viajando, podría interactuar con alguno de los núcleos de hidrógeno (protones), produciendo quizá otras partículas. La predicción de la teoría de la extrañeza es que *nunca* producirá una partícula Λ en una reacción simple como, por ejemplo,

$$K^0 + p \rightarrow \Lambda^0 + \pi^+,$$

aunque un $\bar{K}^0$ puede hacer exactamente eso. Esto es, en una cámara de burbujas, un $\bar{K}^0$ podría producir el evento mostrado en la figura 11-5(b) —en el cual el Λ^0 se puede ver porque se desintegra— pero un K^0 nunca podría. Ésta es la primera parte de nuestra historia: la conservación de la extrañeza.

La conservación de la extrañeza, sin embargo, *no es perfecta*. Hay desintegraciones muy lentas de partículas extrañas —desintegraciones que ocurren en un tiempo largo,‡ algo así como 10^{-10} segundos, en las que la extrañeza *no* se conserva—. Se les llama desintegraciones “débiles”. Por ejemplo, el K^0 se desintegra en un par de mesones π (+ y -) con un tiempo de vida de 10^{-10} segundos. En realidad, fue de este modo que las partículas K fueron observadas por primera vez. Noten que la reacción de desintegración

$$K^0 \rightarrow \pi^+ + \pi^-$$

no conserva la extrañeza, por lo que no puede ir “rápido” mediante una interacción fuerte; solamente puede ocurrir a través del proceso de desintegración débil.

Pero el $\bar{K}^0$ también se desintegra *en la misma forma* —en un π^+ y un π^- — y también con el mismo tiempo de vida

$$\bar{K}^0 \rightarrow \pi^- + \pi^+.$$

De nuevo tenemos una desintegración débil, porque no conserva la extrañeza. Hay un principio según el cual, para cualquier reacción, existe la reacción correspondiente con “materia” remplazada por “antimateria” y *viceversa*. Como $\bar{K}^0$ es la antipartícula de K^0 , debería desintegrarse en las antipartículas de π^+ y π^- , pero la antipartícula de π^+ es π^- . (O *viceversa*, si lo prefieren. Resulta que

† La partícula Λ libre se desintegra muy despacio mediante una interacción *débil* (por lo que no se necesita conservar la extrañeza). Los productos de desintegración son o bien un p y un π^- , o un n y un π^0 . El tiempo de vida es de 2.2×10^{-10} s.

‡ Un tiempo típico para las interacciones fuertes es más o menos de 10^{-23} s.

para los mesones π no importa a cuál llamen “materia”). Así, como consecuencia de las desintegraciones débiles, K^0 y $\bar{K}^0$ pueden dar los mismos productos finales. Cuando “los vemos” a través de sus desintegraciones —en una cámara de burbujas— parecen la misma partícula. Sólo sus interacciones fuertes son diferentes.

Por fin nos encontramos preparados para describir el trabajo de Gell-Mann y Pais. Primero, notaron que debido a que ambos K^0 y $\bar{K}^0$ se pueden transformar en estados de dos mesones π , debe haber alguna amplitud para que un K^0 pueda convertirse en un $\bar{K}^0$, y que también un $\bar{K}^0$ pueda convertirse en un K^0 . Al escribir las reacciones como se hace en química, tendríamos

$$K^0 \rightleftharpoons \pi^- + \pi^+ \rightleftharpoons \bar{K}^0. \quad (11.43)$$

Estas reacciones implican que existe una cierta amplitud por unidad de tiempo, digamos $-i/\hbar$ por $\langle \bar{K}^0 | W | K^0 \rangle$, para que K^0 pase a $\bar{K}^0$ mediante la interacción débil responsable de la desintegración en dos mesones π . Y existe la amplitud correspondiente $\langle K^0 | W | \bar{K}^0 \rangle$ para el proceso inverso. Debido a que la materia y la antimateria se comportan exactamente del mismo modo, estas dos amplitudes son numéricamente iguales; llamaremos a ambas A :

$$\langle \bar{K}^0 | W | K^0 \rangle = \langle K^0 | W | \bar{K}^0 \rangle = A. \quad (11.44)$$

Ahora —dijeron Gell-Mann y Pais— tenemos una situación interesante. Lo que la gente ha estado llamando dos estados distintos del mundo — K^0 y $\bar{K}^0$ — en realidad se debería considerar como *un sistema* de dos estados, porque existe una amplitud para ir de un estado al otro. Para un tratamiento completo, por supuesto, tendríamos que tratar más de dos estados, porque también están los estados de los 2π , y demás; sin embargo, como ellos estaban principalmente interesados en la relación entre K^0 y $\bar{K}^0$, no tuvieron que complicar las cosas y pudieron hacer la aproximación de un sistema de dos estados. Los demás estados *se tomaron* en cuenta en el sentido de que sus efectos aparecieron implícitamente en las amplitudes de la ecuación (11.44).

De acuerdo con esto, Gell-Mann y Pais analizaron la partícula neutra como un sistema de dos estados. Empezaron escogiendo como los dos estados base a los estados $|K^0\rangle$ y $|\bar{K}^0\rangle$. (De ahora en adelante, la historia se parece mucho a la de la molécula de amoníaco.) Cualquier estado $|\psi\rangle$ de la partícula K neutra se podría describir dando las amplitudes para que estuviese en alguno de los estados base. Llamaremos a estas amplitudes

$$C_+ = \langle K^0 | \psi \rangle, \quad C_- = \langle \bar{K}^0 | \psi \rangle. \quad (11.45)$$

El paso siguiente fue escribir las ecuaciones hamiltonianas para este sistema de dos estados. Si no hubiese acoplamiento entre K^0 y $\bar{K}^0$, las ecuaciones serían simplemente

$$\begin{aligned} i\hbar \frac{dC_+}{dt} &= E_0 C_+, \\ i\hbar \frac{dC_-}{dt} &= E_0 C_-. \end{aligned} \quad (11.46)$$

Pero, puesto que existe la amplitud $\langle K^0 | W | \bar{K}^0 \rangle$ para que $\bar{K}^0$ se convierta en K^0 , debe haber el término adicional

$$\langle K^0 | W | \bar{K}^0 \rangle C_- = A C_-$$

sumado al lado derecho de la primera ecuación. Y, del mismo modo, el término $A C_+$ se debería agregar en la ecuación de la rapidez de cambio de C_- .

Pero esto no es todo. Cuando el efecto de dos piones se toma en cuenta, hay una amplitud *adicional* para que K^0 se transforme en *sí mismo*, a través del proceso

$$K^0 \rightarrow \pi^- + \pi^+ \rightarrow K^0.$$

La amplitud adicional, que escribiremos como $\langle K^0 | W | K^0 \rangle$, es igual a la amplitud $\langle \bar{K}^0 | W | K^0 \rangle$, puesto que las amplitudes para ir a y desde un par de mesones π son idénticas para K^0 y $\bar{K}^0$. Si desean, el argumento se puede escribir detalladamente así. Primero escriban[†]

$$\langle \bar{K}^0 | W | K^0 \rangle = \langle \bar{K}^0 | W | 2\pi \rangle \langle 2\pi | W | K^0 \rangle$$

y

$$\langle K^0 | W | K^0 \rangle = \langle K^0 | W | 2\pi \rangle \langle 2\pi | W | K^0 \rangle.$$

Debido a la simetría entre la materia y la antimateria

$$\langle 2\pi | W | K^0 \rangle = \langle 2\pi | W | \bar{K}^0 \rangle,$$

y también

$$\langle K^0 | W | 2\pi \rangle = \langle \bar{K}^0 | W | 2\pi \rangle.$$

De ahí se sigue que $\langle K^0 | W | K^0 \rangle = \langle \bar{K}^0 | W | K^0 \rangle$, y también que $\langle \bar{K}^0 | W | K^0 \rangle = \langle K^0 | W | \bar{K}^0 \rangle$, como dijimos antes. De cualquier modo, hay dos amplitudes adicionales $\langle K^0 | W | K^0 \rangle$ y $\langle \bar{K}^0 | W | \bar{K}^0 \rangle$, ambas iguales a A , que deberían incluirse en las ecuaciones hamiltonianas. La primera da el término AC_+ en el lado derecho de la ecuación para dC_+/dt , y la segunda da el nuevo término AC_- en la ecuación para dC_-/dt . Pensando de este modo, Gell-Mann y Pais concluyeron que las ecuaciones hamiltonianas para el sistema $K^0\bar{K}^0$ deberían ser

$$i\hbar \frac{dC_+}{dt} = E_0 C_+ + AC_- + AC_+, \quad (11.47)$$

$$i\hbar \frac{dC_-}{dt} = E_0 C_- + AC_+ + AC_-.$$

Debemos corregir algo que dijimos en capítulos anteriores: que dos amplitudes como $\langle K^0 | W | \bar{K}^0 \rangle$ y $\langle \bar{K}^0 | W | K^0 \rangle$, que son inversas entre sí, son siempre complejos conjugados. Era verdad cuando hablábamos de partículas que no se desintegran. Pero si se desintegran —y pueden, por lo tanto, “perdersen”— las dos amplitudes no son necesariamente complejos conjugados. Por ello, la igualdad (11.44) no significa que las amplitudes son números reales; son de hecho números complejos. El coeficiente A es, por consiguiente, complejo; y no podemos simplemente incorporarlo a la energía E_0 .

Como habían jugado con frecuencia con espines de electrones y cosas parecidas, nuestros héroes sabían que las ecuaciones hamiltonianas (11.47) significaban que había *otro* par de estados base, que también se podría usar para representar al sistema de la partícula K y que tendría un comportamiento especialmente simple. Dijeron, “sumemos y restemos esas dos ecuaciones. También, midamos todas las energías a partir de E_0 , y usemos unidades para la energía y el tiempo que hagan $\hbar = 1$ ”. (Esto es lo que los físicos teóricos modernos siempre hacen. No cambia el significado físico, pero hace que las ecuaciones tomen una forma más simple.) Así obtuvieron:

$$i \frac{d}{dt} (C_+ + C_-) = 2A(C_+ + C_-), \quad i \frac{d}{dt} (C_+ - C_-) = 0. \quad (11.48)$$

Está claro que las combinaciones de amplitudes $(C_+ + C_-)$ y $(C_+ - C_-)$ actúan independientemente una de otra (corresponden, por supuesto, a los estados estacionarios que estudiamos antes). En consecuencia, concluyeron que sería más conveniente usar una representación diferente para la partícula K . Definieron los dos estados

$$|K_1\rangle = \frac{1}{\sqrt{2}} (|K^0\rangle + |\bar{K}^0\rangle), \quad |K_2\rangle = \frac{1}{\sqrt{2}} (|K^0\rangle - |\bar{K}^0\rangle). \quad (11.49)$$

[†] Aquí estamos haciendo una simplificación. El sistema de las 2π puede tener muchos estados correspondientes a diversos momentum de los mesones π , y deberíamos hacer que el lado derecho de esta ecuación fuera una suma sobre los diversos estados base de los π . El tratamiento completo también nos conduce a las mismas conclusiones.

Dijeron que en lugar de pensar en los mesones K^0 y $\bar{K}^0$, de igual manera podemos pensar en términos de las dos “partículas” (esto es, “estados”) K_1 y K_2 . (Éstos corresponden, por supuesto, a los estados que solíamos llamar $|I\rangle$ y $|II\rangle$). No usamos nuestra notación anterior porque queremos seguir la notación de los autores originales —es la que verán en los seminarios de física—.)

Ahora bien, Gell-Mann y Pais no hicieron todo esto solamente para poner nombres diferentes a las partículas —hay también nueva física extraña en esto—. Supongan que C_1 y C_2 son las amplitudes para que cierto estado $|\psi\rangle$ esté en ya sea un mesón K_1 o un mesón K_2 :

$$C_1 = \langle K_1 | \psi \rangle, \quad C_2 = \langle K_2 | \psi \rangle.$$

De acuerdo con las ecuaciones de (11.49),

$$C_1 = \frac{1}{\sqrt{2}}(C_+ + C_-), \quad C_2 = \frac{1}{\sqrt{2}}(C_+ - C_-). \quad (11.50)$$

Entonces, las ecuaciones (11.48) se convierten en

$$i \frac{dC_1}{dt} = 2AC_1, \quad i \frac{dC_2}{dt} = 0. \quad (11.51)$$

Las soluciones son

$$C_1(t) = C_1(0)e^{-i2At}, \quad C_2(t) = C_2(0), \quad (11.52)$$

donde, naturalmente, $C_1(0)$ y $C_2(0)$ son las amplitudes en $t = 0$.

Estas ecuaciones dicen que si una partícula K neutra empieza en el estado $|K_1\rangle$ en $t = 0$ [de modo que $C_1(0) = 1$ y $C_2(0) = 0$], las amplitudes para el tiempo t son

$$C_1(t) = e^{-i2At}, \quad C_2(t) = 0.$$

Recordando que A es un número complejo, es conveniente tomar $2A = \alpha - i\beta$. (Puesto que la parte imaginaria de $2A$ resulta ser negativa, la escribimos como *menos* $i\beta$.) Con esta sustitución, $C_1(t)$ nos queda

$$C_1(t) = C_1(0)e^{-\beta t}e^{-i\alpha t}. \quad (11.53)$$

La probabilidad para encontrar una partícula K_1 en t es el cuadrado de la norma de esta amplitud, que es $e^{-2\beta t}$. Y, según las ecuaciones (11.52), la probabilidad para encontrar el estado K_2 en cualquier tiempo es cero. Esto significa que, si tienen una partícula K en el estado $|K_1\rangle$, la probabilidad para encontrarla en el mismo estado decrece exponencialmente con el tiempo —pero nunca la encontrarán en el estado $|K_2\rangle$ —. ¿A dónde va? Se desintegra en dos mesones π con vida media $\tau = 1/2\beta$ que es, experimentalmente, 10^{-10} s. Ya preveníamos esto al decir que A era complejo.

Por otra parte, la ecuación (11.52) dice que si tenemos una partícula K completamente en el estado $|K_2\rangle$, permanece así para siempre. Bueno, esto no es realmente verdad. Se ha observado experimentalmente que se desintegra en *tres* mesones π , pero 600 veces más lentamente que la desintegración en dos piones que hemos descrito. Por consiguiente, hay algunos otros términos pequeños que hemos dejado fuera en nuestra aproximación. Pero siempre que consideremos solamente la desintegración en dos piones, el K_2 dura “para siempre”.

Para terminar con la historia de Gell-Mann y Pais. Después consideraron lo que sucede cuando se produce una partícula K a partir de una partícula Λ^0 en una interacción fuerte. Como debe tener extrañeza $+1$, se debe producir en el estado K^0 . Por lo tanto, para $t = 0$ no es un K_1 ni un K_2 sino una *mezcla*. Las condiciones iniciales son

$$C_+(0) = 1, \quad C_-(0) = 0.$$

Pero esto significa —de acuerdo con la ecuación (11.50)— que

$$C_1(0) = \frac{1}{\sqrt{2}}, \quad C_2(0) = \frac{1}{\sqrt{2}},$$

y —de acuerdo con (11.52) y con (11.53)— que

$$C_1(t) = \frac{1}{\sqrt{2}} e^{-\beta t} e^{-i\alpha t}, \quad C_2(t) = \frac{1}{\sqrt{2}}. \quad (11.54)$$

Ahora recuerden que K^0 y $\bar{K}^0$ son combinaciones lineales de K_1 y K_2 . En las ecuaciones (11.54), las amplitudes se escogieron de forma que en $t = 0$ las partes $\bar{K}^0$ se cancelen entre sí por interferencia, dejando solamente un estado K^0 . Pero el estado $|K_1\rangle$ *varía en el tiempo*, y el estado $|K_2\rangle$ *no lo hace*. Después de $t = 0$ la interferencia de C_1 y C_2 dará amplitudes finitas para ambos K^0 y $\bar{K}^0$.

¿Qué significa todo esto? Volvamos y consideremos el experimento que diseñamos en la figura 11-5. Un mesón π^- ha producido una partícula Λ^0 y un mesón K^0 , el cual anda sigilosamente por el hidrógeno de la cámara. Conforme avanza, hay cierta probabilidad pequeña, pero uniforme, para que colisione con un núcleo de hidrógeno. Antes pensábamos que la conservación de la extrañeza impediría que la partícula K produjera una Λ^0 en tal interacción. Ahora, sin embargo, vemos que no es así. Porque, aunque nuestra partícula K *empiece* como un K^0 —que no puede darnos una Λ^0 — no se *queda* así. Después de algún tiempo, *hay cierta amplitud* para que haya cambiado al estado $\bar{K}^0$. Podemos, por lo tanto, esperar ver algunas veces una Λ^0 que se ha producido en el camino de la partícula K . La probabilidad para que esto ocurra está dada por la amplitud C_- , que podemos relacionar [usando las ecuaciones (11.50) a la inversa] con C_1 y C_2 . La relación es

$$C_- = \frac{1}{\sqrt{2}} (C_1 - C_2) = \frac{1}{2} (e^{-\beta t} e^{-i\alpha t} - 1). \quad (11.55)$$

Cuando nuestra partícula K avance, la probabilidad para que “actúe como” un $\bar{K}^0$ es igual a $|C_-|^2$, que es

$$|C_-|^2 = \frac{1}{4} (1 + e^{-2\beta t} - 2e^{-\beta t} \cos \alpha t). \quad (11.56)$$

¡Un resultado extraño y complicado!

Ésta es entonces la notable predicción de Gell-Mann y Pais: cuando se produce un K^0 , la probabilidad para que se convierta en un $\bar{K}^0$ —como puede demostrar siendo capaz de producir una Λ^0 — varía con el tiempo de acuerdo a la ecuación (11.56). Esta predicción surgió usando lógica pura y los principios básicos de la mecánica cuántica —sin ninguna clase de conocimiento de los procesos internos de la partícula K —. Como nadie sabe nada acerca del mecanismo interno, hasta aquí es hasta donde Gell-Mann y Pais pudieron avanzar. No pudieron dar ningún valor teórico para α y β . Y nadie ha podido hacerlo hasta la fecha. Ellos pudieron dar un valor para β obtenido a partir de la rapidez de desintegración en dos π observada experimentalmente ($2\beta = 10^{10} \text{ s}^{-1}$), pero no pudieron decir nada acerca de α .

Hemos graficado la función de la ecuación (11.56) para dos valores de α en la figura 11-6. Pueden ver que la forma depende mucho del cociente entre α y β . No hay probabilidad para $\bar{K}^0$ al principio; se forma después. Si α es grande, la probabilidad tendría grandes oscilaciones. Si α es pequeña, habrá muy pocas o ninguna oscilación —la probabilidad sólo aumentará suavemente hasta $1/4$ —.

Ahora bien, típicamente la partícula K estará viajando a una velocidad constante muy próxima a la velocidad de la luz. Las curvas de la figura 11-6 también representan la probabilidad de observar un $\bar{K}^0$ a lo largo del trayecto —con distancias típicas de varios centímetros—. Pueden ver por qué esta predicción es tan notablemente peculiar. Producen ustedes una sola partícula y en lugar de sólo desintegrarse, hace algo más. Algunas veces se desintegra y otras se convierte en otra clase de partícula. Su probabilidad característica para producir un efecto varía en una forma rara según avanza. No hay cosa que se le parezca en la naturaleza. Y esta predicción tan notable se hizo únicamente mediante razonamientos acerca de la interferencia de amplitudes.

Si hay algún lugar donde tengamos la posibilidad de comprobar los principios generales de la mecánica cuántica del modo más puro —¿funciona o no la

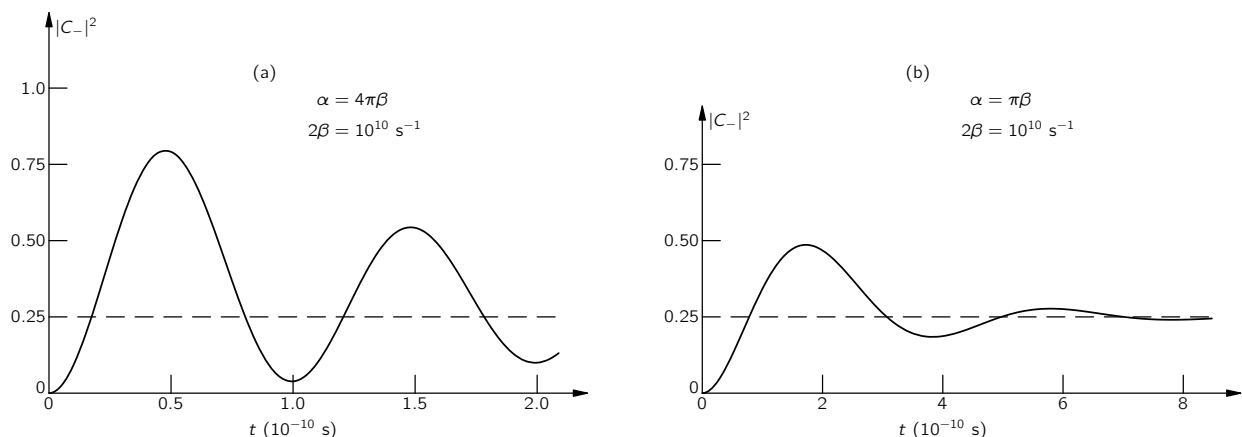


Figura 11-6. La función de la ecuación (11.56): (a) para $\alpha = 4\pi\beta$, (b) para $\alpha = \pi\beta$ (con $2\beta = 10^{10} \text{ s}^{-1}$).

superposición de amplitudes?— es éste. A pesar de que este efecto se predijo hace algunos años, no hay determinación experimental muy clara. Hay algunos resultados vagos que indican que α no es cero, y que el efecto ocurre en realidad —éstos indican que α se encuentra entre 2β y 4β —. Esto es todo lo que hay, de manera experimental. Sería magnífico comprobar exactamente la curva, para ver si el principio de superposición sirve aún en un mundo tan misterioso como el de las partículas extrañas —con razones desconocidas para las desintegraciones y razones desconocidas para la extrañeza—.

El análisis que acabamos de hacer es muy característico del modo en que se usa la mecánica cuántica hoy en día en la búsqueda de una comprensión de las partículas extrañas. Todas las complicadas teorías de las que puedan oír hablar, no son ni más ni menos que esta clase de artificios elementales que utilizan los principios de superposición, y otros principios de la mecánica cuántica que están en el mismo nivel. Algunos dicen tener teorías por las que es posible calcular α y β , o por lo menos α dada β , pero estas teorías son completamente inútiles. Por ejemplo, la teoría que predice el valor de α dado el de β , nos dice que el valor de α debería ser infinito. El conjunto de ecuaciones con el que originalmente inician incluye a los dos mesones π , y después va de los dos π de regreso a K^0 , y así sucesivamente. Una vez desarrollado todo esto, de verdad produce un par de ecuaciones como las que nosotros tenemos; pero como hay un número infinito de estados de los dos π que dependen de su momentum, al integrar sobre todas las posibilidades se obtiene un α que es infinito. Pero el α de la naturaleza *no* es infinito. Por lo tanto, las teorías dinámicas son erróneas. Es en verdad muy notable el que todos los fenómenos que *realmente* se puede predecir en el mundo de las partículas extrañas surgen de los principios de la mecánica cuántica al nivel que ustedes están aprendiendo ahora.

11-6 Generalización a sistemas de N estados

Hemos terminado con todos los sistemas de dos estados de que queríamos hablar. En los capítulos siguientes continuaremos estudiando sistemas con más estados. La extensión a sistemas de N estados de las ideas que hemos elaborado para los sistemas de dos estados es bastante directa. Se hace así.

Si un sistema tiene N estados distintos, podemos representar cualquier estado $|\psi(t)\rangle$ como combinación lineal de cualquier conjunto de estados base $|i\rangle$, donde $i = 1, 2, 3, \dots, N$;

$$|\psi(t)\rangle = \sum_{\text{toda } i} |i\rangle C_i(t). \quad (11.57)$$

Los coeficientes $C_i(t)$ son las amplitudes $\langle i | \psi(t) \rangle$. El comportamiento de las amplitudes C_i en el tiempo está regido por las ecuaciones

$$i\hbar \frac{dC_i(t)}{dt} = \sum_j H_{ij} C_j, \quad (11.58)$$

donde la matriz de energía H_{ij} describe la física del problema. Se ve igual que para dos estados. Solamente que ahora, ambas i y j deben recorrer los N estados base, y la matriz de energía H_{ij} —o, si lo prefieren, el hamiltoniano— es una matriz de $N \times N$ con N^2 números. Como antes, $H_{ij}^* = H_{ji}$ —siempre que las partículas se conserven— y los elementos de la diagonal H_{ii} son números reales.

Encontramos una solución general para las C de un sistema de dos estados cuando la matriz de energía es constante (no depende de t). Tampoco es difícil resolver la ecuación (11.58) para un sistema con N estados cuando H no depende del tiempo. De nuevo, empezamos buscando una solución posible en la que todas las amplitudes tengan la *misma* dependencia temporal. Probamos

$$C_i = a_i e^{-(i/\hbar)Et}. \quad (11.59)$$

Al sustituir estas C_i en (11.58), las derivadas $dC_i(t)/dt$ se convierten simplemente en $(-i/\hbar)EC_i$. Cancelando el factor exponencial común en todos los términos, obtenemos

$$Ea_i = \sum_j H_{ij} a_j. \quad (11.60)$$

Éste es un conjunto de N ecuaciones algebraicas lineales para las N incógnitas $a_1, a_2, \dots, a_N$, y existe una solución sólo si tienen suerte —solamente si el determinante de los coeficientes de todas las a es cero—. Pero no es necesario ser tan sutiles; pueden empezar simplemente a resolver las ecuaciones en la forma que quieran, y encontrarán que únicamente se pueden resolver para ciertos valores de E . (Recuerden que E es lo único ajustable que tenemos en estas ecuaciones.)

Sin embargo, si quieren ser formales, pueden escribir la ecuación (11.60) en la forma

$$\sum_j (H_{ij} - \delta_{ij}E) a_j = 0. \quad (11.61)$$

Después pueden usar la regla —si la conocen— de que estas ecuaciones solamente tendrán una solución para aquellos valores de E para los cuales

$$\text{Det}(H_{ij} - \delta_{ij}E) = 0. \quad (11.62)$$

Cada término del determinante es sólo H_{ij} , excepto E se resta de todos los elementos diagonales. Esto es, (11.62) es exactamente

$$\text{Det} \begin{pmatrix} H_{11} - E & H_{12} & H_{13} & \dots \\ H_{21} & H_{22} - E & H_{23} & \dots \\ H_{31} & H_{32} & H_{33} - E & \dots \\ \dots & \dots & \dots & \dots \end{pmatrix} = 0. \quad (11.63)$$

Éste es, naturalmente, sólo un modo especial de escribir una ecuación algebraica para E , que es la suma de un grupo de productos de todos los términos tomados de un cierto modo. Estos productos darán todas las potencias de E hasta E^N .

Entonces tenemos un polinomio de grado N igual a cero y hay, en general, N raíces. (Debemos recordar, no obstante, que algunas de ellas pueden ser raíces con multiplicidad —lo que significa que dos o más son iguales—.) Llamemos a las N raíces

$$E_I, E_{II}, E_{III}, \dots, E_{\mathbf{n}}, \dots, E_{\mathbf{N}}. \quad (11.64)$$

(Usaremos $\mathbf{n}$ para representar el n -ésimo número romano, por lo que $\mathbf{n}$ toma los valores $I, II, \dots, \mathbf{N}$.) Puede ocurrir que algunas de estas energías sean iguales —digamos $E_{II} = E_{III}$ —, pero aun así las llamaremos con nombres diferentes.

Las ecuaciones (11.60) —u (11.61)— tienen una solución para cada valor de E . Si sustituyen cualquiera de las E —por ejemplo $E_{\mathbf{n}}$ — en (11.60) y despejan a_i , obtienen un conjunto que pertenece a la energía $E_{\mathbf{n}}$. Llamaremos a este conjunto $a_i(\mathbf{n})$.

Usando estas $a_i(\mathbf{n})$ en (11.59) tenemos las amplitudes $C_i(\mathbf{n})$ para que los estados de energía definida estén en los estados base $|i\rangle$. Si $|\mathbf{n}\rangle$ representa al vector de estado para el estado de energía definida en $t = 0$, podemos escribir

$$C_i(\mathbf{n}) = \langle i | \mathbf{n} \rangle e^{-(i/\hbar)E_{\mathbf{n}}t},$$

con

$$\langle i | \mathbf{n} \rangle = a_i(\mathbf{n}). \quad (11.65)$$

El estado completo de energía definida $|\psi_{\mathbf{n}}(t)\rangle$ se puede escribir entonces en la forma

$$|\psi_{\mathbf{n}}(t)\rangle = \sum_i |i\rangle a_i(\mathbf{n}) e^{-(i/\hbar)E_{\mathbf{n}}t},$$

o

$$|\psi_{\mathbf{n}}(t)\rangle = |\mathbf{n}\rangle e^{-(i/\hbar)E_{\mathbf{n}}t}. \quad (11.66)$$

Los vectores de estado $|\mathbf{n}\rangle$ describen la configuración de los estados de energía definida, pero no tiene los factores de la dependencia temporal. Por eso son vectores constantes, que se pueden emplear como un nuevo conjunto base si lo deseamos.

Cada uno de los estados $|\mathbf{n}\rangle$ tiene la propiedad —como pueden demostrar fácilmente— de que cuando los aplicamos al operador hamiltoniano $\hat{H}$, nos da exactamente $E_{\mathbf{n}}$ por el mismo estado:

$$\hat{H} |\mathbf{n}\rangle = E_{\mathbf{n}} |\mathbf{n}\rangle. \quad (11.67)$$

La energía $E_{\mathbf{n}}$ es, entonces, un número característico del operador hamiltoniano $\hat{H}$. Como ya hemos visto, un hamiltoniano tendrá, en general, varias energías características. En el mundo de las matemáticas se les llama “valores propios” de la matriz H_{ij} . Los físicos los llaman “eigenvalores” de $\hat{H}$. (“Eigen” es la palabra alemana para “característico” o “propio”.) Para cada eigenvalor de $\hat{H}$ —en otras palabras, para cada energía— hay un estado de energía definida, que hemos llamado “estado estacionario”. Los físicos por lo general llaman a los estados $|\mathbf{n}\rangle$ “los estados propios de $\hat{H}$ ”. Cada estado propio corresponde a un eigenvalor particular $E_{\mathbf{n}}$.

Por lo general, los estados $|\mathbf{n}\rangle$ —de los que hay N — también se pueden usar como conjunto base. Para que sea así, todos los estados deben ser ortogonales, lo que quiere decir que para cualquier par de ellos, digamos $|\mathbf{n}\rangle$ y $|\mathbf{m}\rangle$,

$$\langle \mathbf{n} | \mathbf{m} \rangle = 0. \quad (11.68)$$

Esto será automáticamente verdad si todas las energías son diferentes. También, podemos multiplicar todas las $a_i(\mathbf{n})$ por un factor adecuado para que todos los estados estén normalizados —con lo cual queremos decir

$$\langle \mathbf{n} | \mathbf{n} \rangle = 1 \quad (11.69)$$

para todo $\mathbf{n}$ —.

Cuando suceda que la ecuación (11.63) tenga accidentalmente dos (o más) raíces con la misma energía, hay complicaciones menores. En primer lugar, todavía hay dos conjuntos diferentes de a_i que acompañan a las dos energías iguales, pero los estados que se obtienen pueden *no* ser ortogonales. Supongan que siguen el procedimiento normal y encuentran dos estados estacionarios con las mismas energías —llamémosles $|\mu\rangle$ y $|\nu\rangle$ —. Entonces no necesariamente resultarán ortogonales —si no tienen suerte,

$$\langle \mu | \nu \rangle \neq 0—.$$

Sin embargo, es verdad que siempre pueden construir dos estados nuevos, a los que llamaremos $|\mu'\rangle$ y $|\nu'\rangle$, que tengan la misma energía y que sean también ortogonales, de modo que

$$\langle\mu'|\nu'\rangle = 0. \quad (11.70)$$

Pueden hacer esto construyendo a $|\mu'\rangle$ y $|\nu'\rangle$ por medio de una combinación lineal conveniente de $|\mu\rangle$ y $|\nu\rangle$, escogiendo los coeficientes de tal modo que cumplan con la ecuación (11.70). Siempre es conveniente hacer eso. Por lo general supondremos que esto ha sido hecho y que, por lo tanto, siempre podemos asumir que todos nuestros estados de energía propia $|\mathbf{n}\rangle$ son ortogonales.

Nos gustaría, por diversión, probar que cuando dos de los estados estacionarios tienen energías diferentes, son en realidad ortogonales. Para el estado $|\mathbf{n}\rangle$ con la energía $E_{\mathbf{n}}$, tenemos que

$$\hat{H}|\mathbf{n}\rangle = E_{\mathbf{n}}|\mathbf{n}\rangle. \quad (11.71)$$

Esta ecuación operacional en realidad significa que hay una ecuación entre números. Completando las partes que faltan, significa lo mismo que

$$\sum_j \langle i|\hat{H}|j\rangle\langle j|\mathbf{n}\rangle = E_{\mathbf{n}}\langle i|\mathbf{n}\rangle. \quad (11.72)$$

Si tomamos el complejo conjugado de esta ecuación, obtenemos

$$\sum_j \langle i|\hat{H}|j\rangle^*\langle j|\mathbf{n}\rangle^* = E_{\mathbf{n}}^*\langle i|\mathbf{n}\rangle^*. \quad (11.73)$$

Recuerden que el complejo conjugado de una amplitud es la amplitud inversa, por lo que (11.73) se puede escribir de nuevo como

$$\sum_j \langle \mathbf{n}|j\rangle\langle j|\hat{H}|i\rangle = E_{\mathbf{n}}^*\langle \mathbf{n}|i\rangle. \quad (11.74)$$

Ya que esta ecuación es válida para *cualquier* i , su forma “abreviada” es

$$\langle \mathbf{n}|\hat{H} = E_{\mathbf{n}}^*\langle \mathbf{n}|, \quad (11.75)$$

que se conoce como la *adjunta* de la ecuación (11.71).

Ahora podemos probar fácilmente que $E_{\mathbf{n}}$ es un número real. Multiplicamos la ecuación (11.71) por $\langle \mathbf{n}|$ para obtener

$$\langle \mathbf{n}|\hat{H}|\mathbf{n}\rangle = E_{\mathbf{n}}, \quad (11.76)$$

puesto que $\langle \mathbf{n}|\mathbf{n}\rangle = 1$. Entonces multiplicamos a la ecuación (11.75) por la derecha por $|\mathbf{n}\rangle$ y obtenemos

$$\langle \mathbf{n}|\hat{H}|\mathbf{n}\rangle = E_{\mathbf{n}}^*. \quad (11.77)$$

Al comparar (11.76) con (11.77) es claro que

$$E_{\mathbf{n}} = E_{\mathbf{n}}^*, \quad (11.78)$$

lo que significa que $E_{\mathbf{n}}$ es real. Podríamos borrar el asterisco de $E_{\mathbf{n}}$ en la ecuación (11.75).

Finalmente, estamos preparados para demostrar que los estados de energía diferente son ortogonales. Supongamos que $|\mathbf{n}\rangle$ y $|\mathbf{m}\rangle$ son cualquier par de estados base de energía definida. Usando la ecuación (11.75) para el estado $\mathbf{m}$, y multiplicándola por $|\mathbf{n}\rangle$, obtenemos

$$\langle \mathbf{m}|\hat{H}|\mathbf{n}\rangle = E_{\mathbf{m}}\langle \mathbf{m}|\mathbf{n}\rangle.$$

Pero si multiplicamos (11.71) por $\langle \mathbf{m}|$, obtenemos

$$\langle \mathbf{m}|\hat{H}|\mathbf{n}\rangle = E_{\mathbf{n}}\langle \mathbf{m}|\mathbf{n}\rangle.$$

Ya que los términos en los lados izquierdos de estas dos ecuaciones son iguales, los términos en los lados derechos lo son también, así:

$$E_{\mathbf{m}}\langle \mathbf{m} | \mathbf{n} \rangle = E_{\mathbf{n}}\langle \mathbf{m} | \mathbf{n} \rangle. \quad (11.79)$$

Si $E_{\mathbf{m}} = E_{\mathbf{n}}$, la ecuación no nos dice nada. Pero si las energías de los dos estados $|\mathbf{m}\rangle$ y $|\mathbf{n}\rangle$ *son diferentes* ($E_{\mathbf{m}} \neq E_{\mathbf{n}}$), la ecuación (11.79) nos dice que $\langle \mathbf{m} | \mathbf{n} \rangle$ debe ser cero, como queríamos demostrar. Los dos estados son necesariamente ortogonales siempre que $E_{\mathbf{n}}$ y $E_{\mathbf{m}}$ sean numéricamente diferentes.

La estructura hiperfina en el hidrógeno

12-1 Estados base para un sistema de dos partículas con espín un medio

En este capítulo trataremos la “separación hiperfina” del hidrógeno, porque es un ejemplo interesante, en sentido físico, de lo que ya podemos hacer con la mecánica cuántica. Es un ejemplo con más de dos estados, e ilustrará cómo se aplican los métodos de la mecánica cuántica a problemas un poco más complicados. Es lo suficientemente más complicado como para que, una vez que hayan visto cómo se lo trata, puedan obtener inmediatamente la generalización para todo tipo de problemas.

Como saben, el átomo de hidrógeno consiste en un electrón que está en los alrededores del protón, donde puede existir en cualquiera de los estados de energía discretos, para cada uno de los cuales el patrón del movimiento del electrón es diferente. Por ejemplo, el primer estado excitado está a $3/4$ de Rydberg, o sea unos 10 electronvolts, por encima del estado fundamental. Pero aun lo que se denomina estado fundamental del hidrógeno no es realmente un solo estado de energía definida, a causa de los espines del electrón y el protón. Estos espines son responsables de la “estructura hiperfina” de los niveles de energía, la cual separa todos los niveles de energía en varios niveles casi iguales.

El electrón puede tener su espín “hacia arriba” o “hacia abajo”, y el protón también puede tener su espín “hacia arriba” o “hacia abajo”. Hay, en consecuencia, *cuatro* estados posibles para el espín por cada estado dinámico del átomo. Es decir, cuando la gente dice “el estado fundamental” del hidrógeno, realmente quieren decir los “cuatro estados fundamentales”, y no simplemente el estado más bajo. Los cuatro estados de espín no tienen exactamente la misma energía; hay pequeñas variaciones respecto a las energías que esperaríamos si no hubiera espines. Sin embargo, las variaciones son mucho menores que los 10 electronvolts que hay aproximadamente entre el estado fundamental y el que le sigue. En consecuencia, cada estado dinámico tiene su energía separada en un conjunto de niveles de energía muy cercanos —la llamada *separación hiperfina*—.

En este capítulo queremos calcular las diferencias de energía entre los cuatro estados de espín. La separación hiperfina se debe a la interacción de los momentos magnéticos del electrón y el protón, la cual da una energía magnética ligeramente diferente para cada estado de espín. Estas variaciones de energía son del orden de diezmillonésimos de un electronvolt —¡realmente muy pequeños comparados con 10 volts!—. Es por este gran espaciamiento de energía que podemos considerar el estado fundamental del hidrógeno como un sistema de “cuatro estados”, sin preocuparnos realmente de que haya muchos otros estados a energías más altas. Aquí nos limitaremos a un estudio de la estructura hiperfina del estado fundamental del átomo de hidrógeno.

Para nuestros propósitos, no nos interesa ninguno de los detalles sobre las *posiciones* del electrón y del protón, porque todo eso ha sido resuelto por el átomo, por así decirlo —él mismo lo resolvió al ir al estado fundamental—. Sólo necesitamos saber que tenemos un electrón y un protón cercanos en alguna relación espacial determinada. Además, sus espines pueden tener varias orientaciones relativas diferentes. Es sólo el efecto de los espines lo que queremos examinar.

La primera pregunta que tenemos que contestar es: ¿cuáles son los *estados base* del sistema? En realidad, la pregunta está formulada de manera incorrecta. No hay una cosa como “*los*” estados base, porque, naturalmente, el conjunto de estados base que elijan no es único. Siempre se puede obtener nuevos conjuntos mediante combinaciones lineales de los estados originales. Siempre hay muchas

12-1 Estados base para un sistema de dos partículas con espín un medio

12-2 El hamiltoniano para el estado fundamental del hidrógeno

12-3 Los niveles de energía

12-4 La separación de Zeeman

12-5 Los estados en un campo magnético

12-6 La matriz de proyección para espín uno

posibilidades para escoger los estados base, y todas ellas son igualmente válidas. De ahí que la pregunta no sea ¿cuál es *el* conjunto base?, sino ¿cuál *podría* ser un conjunto base? Podemos elegir el que más nos convenga. Por lo general es mejor comenzar con un conjunto base que sea el más claro *físicamente*. Puede que no sea la solución para ningún problema, o que no tenga ninguna importancia *directa*, pero por lo general servirá para que sea más fácil comprender lo que pasa.

Elegimos los cuatro estados base siguientes:

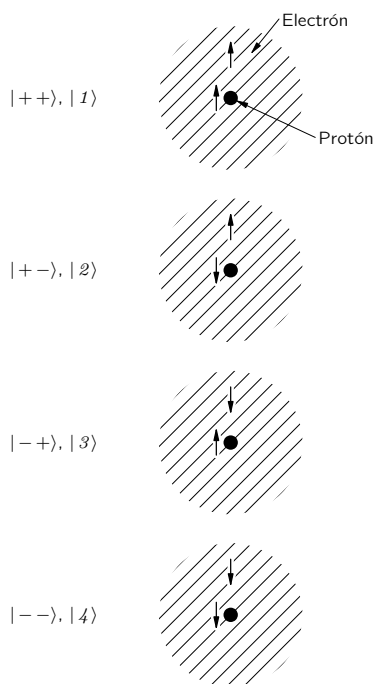
Estado 1: El electrón y el protón están ambos “hacia arriba”.

Estado 2: El electrón está “hacia arriba” y el protón “hacia abajo”.

Estado 3: El electrón está “hacia abajo” y el protón “hacia arriba”.

Estado 4: El electrón y el protón están ambos “hacia abajo”.

Necesitamos una notación práctica para estos estados, por lo que los representaremos como sigue:



Estado 1: $|++\rangle$; electrón *hacia arriba*, protón *hacia arriba*.

Estado 2: $|+-\rangle$; electrón *hacia arriba*, protón *hacia abajo*.

Estado 3: $|-+\rangle$; electrón *hacia abajo*, protón *hacia arriba*.

Estado 4: $|--\rangle$; electrón *hacia abajo*, protón *hacia abajo*.

(12.1)

Tendrán que recordar que el *primer* signo más o menos se refiere al electrón y el *segundo* al protón. Para tener una referencia cómoda, hemos resumido esta notación en la figura 12-1. A veces resultará conveniente llamar a estos estados $|1\rangle$, $|2\rangle$, $|3\rangle$ y $|4\rangle$.

Ustedes podrán decir: “Pero las partículas interactúan, y puede que éstos no sean los estados base correctos. Parece como si estuviera considerando a las dos partículas independientemente”. ¡Claro que sí! La interacción provoca el problema: cuál sea el *hamiltoniano* para el sistema, pero la interacción no interviene en la pregunta de cómo *describir* al sistema. Lo que elijamos como estados base no tiene nada que ver con lo que ocurre después. Es posible que el átomo no pueda *quedarse* en uno de esos estados base, aunque haya empezado en éste. Ésa es otra cuestión. Esa pregunta es: ¿cómo cambian las amplitudes en el tiempo en una base determinada (fija)? Al elegir los estados base, sólo estamos eligiendo los “vectores unitarios” para nuestra descripción.

Mientras estamos en el tema, examinemos el problema general de encontrar un conjunto de estados base cuando hay más de una partícula. Ustedes conocen los estados base para una sola partícula. Un electrón, por ejemplo, está descrito completamente en la vida real —no en nuestros casos simplificados, sino en la vida real— dando las amplitudes para que esté en cada uno de los siguientes estados:

$|\text{electrón “hacia arriba” con momentum } \mathbf{p}\rangle$

o

$|\text{electrón “hacia abajo” con momentum } \mathbf{p}\rangle.$

En realidad hay dos conjuntos infinitos de estados, un estado para cada valor de $\mathbf{p}$. Es decir, un estado del electrón $|\psi\rangle$ queda descrito completamente, si conocen todas las amplitudes

$$\langle +, \mathbf{p} | \psi \rangle \quad \text{y} \quad \langle -, \mathbf{p} | \psi \rangle,$$

donde los $+$ y $-$ representan las componentes del momentum angular en la dirección de algún eje —usualmente el eje z — y $\mathbf{p}$ es el vector momentum. Debe haber, por consiguiente, dos amplitudes para cada momentum posible (un conjunto multi-infinito de estados base). Eso es todo lo que se necesita para describir una sola partícula.

Cuando hay más de una partícula, los estados base se pueden escribir en forma semejante. Por ejemplo, si hubiera un electrón y un protón en una situación más

Figura 12-1. Un conjunto de estados base para el estado fundamental del átomo de hidrógeno.

complicada que la que estamos considerando, los estados base podrían ser del siguiente tipo:

|un electrón con espín “hacia arriba”, moviéndose con momentum $\mathbf{p}_1$ y
un protón con espín “hacia abajo”, moviéndose con momentum $\mathbf{p}_2$ ⟩.

Y así sucesivamente para otras combinaciones del espín. Si hay más de dos partículas —la misma idea—. Como pueden ver, escribir los estados base *posibles* es realmente muy fácil. El único problema es: ¿cuál es el hamiltoniano?

Para estudiar el estado fundamental del hidrógeno, no es necesario que utilizemos los conjuntos completos de estados base para los diversos valores del momentum. Especificamos estados determinados de momentum para el protón y el electrón cuando decimos “el estado fundamental”. Es posible calcular los detalles de la configuración —las amplitudes para todos los estados base del momentum— pero eso es otro problema. En este momento sólo nos interesan los efectos del espín, por lo que sólo tomamos los cuatro estados base de (12.1). El siguiente problema es: ¿cuál es el hamiltoniano para este conjunto de estados?

12-2 El hamiltoniano para el estado fundamental del hidrógeno

Dentro de un momento les diremos cuál es. Primero tenemos que recordarles una cosa: *cualquier* estado siempre se puede escribir como combinación de los estados base. Para cualquier estado $|\psi\rangle$ podemos escribir:

$$|\psi\rangle = |++\rangle\langle ++|\psi\rangle + |+-\rangle\langle +-|\psi\rangle + |-+\rangle\langle -+|\psi\rangle + |--\rangle\langle --|\psi\rangle. \quad (12.2)$$

Recuerden que los corchetes completos son sólo números complejos, por lo cual podemos escribirlos de la manera habitual como C_i , donde $i = 1, 2, 3$ o 4 , y escribir la ecuación (12.2) en la forma

$$|\psi\rangle = |++\rangle C_1 + |+-\rangle C_2 + |-+\rangle C_3 + |--\rangle C_4. \quad (12.3)$$

Al dar las cuatro amplitudes C_i describimos completamente el estado de espín $|\psi\rangle$. Si estas cuatro amplitudes cambian en el tiempo, como lo harán, la rapidez de variación temporal está dada por el operador $\hat{H}$. El problema es encontrar $\hat{H}$.

No hay una regla general para escribir el hamiltoniano de un sistema atómico, y encontrar la fórmula apropiada requiere mucha más habilidad que encontrar un conjunto de estados base. Pudimos darles una regla general para escribir un conjunto de estados base para cualquier problema con un protón y un electrón, pero describir el hamiltoniano general de esa combinación es muy difícil a este nivel. En vez de ello, los conduciremos a un hamiltoniano por un argumento heurístico —y tendrán que aceptarlo como correcto, porque los resultados estarán de acuerdo con la observación experimental—.

Recordarán ustedes que en el capítulo anterior pudimos describir el hamiltoniano de una sola partícula de espín un medio, utilizando las matrices sigma —o los exactamente equivalentes operadores sigma—. La tabla 12-1 resume las propiedades de los operadores. Estos operadores —que son simplemente una manera abreviada y conveniente de seguirle la pista a elementos matriciales del tipo $\langle +|\sigma_z|+\rangle$ — fueron útiles para describir el comportamiento de una *sola* partícula con espín un medio. La pregunta es: ¿podemos encontrar un mecanismo análogo para describir un sistema con dos espines? La respuesta es sí, de manera sencilla, como sigue. Inventamos algo que llamaremos “sigma electrón”, que representaremos con el operador vectorial σ^e y que tiene como componentes x , y y z a σ_x^e , σ_y^e y σ_z^e . Ahora adoptamos la *convención* de que cuando una de estas cosas opera sobre cualquiera de nuestros estados base del átomo de hidrógeno, solamente actúa sobre el espín del *electrón*, y exactamente de la misma manera como si el electrón estuviera solo. Por ejemplo: ¿cuánto vale $\sigma_y^e|-\rangle$? Como

Tabla 12-1

$\sigma_z +\rangle = + +\rangle$
$\sigma_z -\rangle = - -\rangle$
$\sigma_x +\rangle = + -\rangle$
$\sigma_x -\rangle = + +\rangle$
$\sigma_y +\rangle = +i -\rangle$
$\sigma_y -\rangle = -i +\rangle$

σ_y en un electrón “hacia abajo” es $-i$ por el estado correspondiente al electrón “hacia arriba”,

$$\sigma_y^e | - + \rangle = -i | + + \rangle.$$

(Cuando σ_y^e actúa sobre el estado combinado, voltea al electrón, pero no hace nada al protón y multiplica el resultado por $-i$.) Operando sobre los otros estados, entonces σ_y^e da

$$\sigma_y^e | + + \rangle = i | - + \rangle,$$

$$\sigma_y^e | + - \rangle = i | - - \rangle,$$

$$\sigma_y^e | - - \rangle = -i | + - \rangle.$$

Sólo recuerden que los operadores σ^e actúan únicamente sobre el *primer* símbolo de espín —o sea, sobre el espín del *electrón*—.

Después definimos el correspondiente operador “sigma protón” para el espín del protón. Sus tres componentes σ_x^p , σ_y^p , σ_z^p actúan del mismo modo que las de σ^e , sólo que sobre el espín del *protón*. Por ejemplo, si tenemos a σ_x^p actuando sobre cada uno de los cuatro estados base, obtenemos —utilizando siempre la tabla 12-1—

$$\sigma_x^p | + + \rangle = | + - \rangle,$$

$$\sigma_x^p | + - \rangle = | + + \rangle,$$

$$\sigma_x^p | - + \rangle = | - - \rangle,$$

$$\sigma_x^p | - - \rangle = | - + \rangle.$$

Como ven, no es muy difícil.

Ahora bien, en el caso más general, podríamos tener cuestiones más complejas. Por ejemplo, podríamos tener productos de dos operadores, como $\sigma_y^e \sigma_z^p$. Si tenemos un producto como éste, hacemos primero lo que dice el operador de la derecha y luego lo que dice el otro.† Por ejemplo, podríamos tener

$$\sigma_x^e \sigma_z^p | + - \rangle = \sigma_x^e (\sigma_z^p | + - \rangle) = \sigma_x^e (- | + - \rangle) = -\sigma_x^e | + - \rangle = - | - - \rangle.$$

Observen que estos operadores no hacen nada sobre números puros —hemos usado esto al escribir $\sigma_x^e(-1) = (-1)\sigma_x^e$ —. Decimos que los operadores “conmutan” con números puros, o sea que un número “se puede pasar al otro lado” del operador. Pueden practicar demostrando que el producto $\sigma_x^e \sigma_z^p$ da los siguientes resultados para los cuatro estados:

$$\sigma_x^e \sigma_z^p | + + \rangle = + | - + \rangle,$$

$$\sigma_x^e \sigma_z^p | + - \rangle = - | - - \rangle,$$

$$\sigma_x^e \sigma_z^p | - + \rangle = + | + + \rangle,$$

$$\sigma_x^e \sigma_z^p | - - \rangle = - | + - \rangle.$$

Si tomamos todos los operadores posibles, usando cada tipo de operador sólo una vez, hay dieciséis posibilidades. Sí, *dieciséis* —siempre que también incluyamos al “operador unidad” $\hat{1}$ —. En primer lugar, tenemos los tres: σ_x^e , σ_y^e , σ_z^e . Luego están los tres σ_x^p , σ_y^p , σ_z^p —seis en total—. Además, están los nueve productos posibles de la forma $\sigma_x^e \sigma_y^p$, lo cual forma un total de 15. Y está el operador unidad, que sencillamente no cambia ningún estado. Dieciséis en total.

Observen ahora que, para un sistema de cuatro estados, la matriz hamiltoniana tiene que ser una matriz de coeficientes de cuatro por cuatro —tendrá dieciséis elementos—. Se puede demostrar fácilmente que cualquier matriz de cuatro por cuatro —y, por lo tanto, la matriz hamiltoniana en particular— se puede expresar como combinación lineal de las dieciséis matrices de doble espín correspondientes

† Para estos operadores *en particular*, notarán que no importa el orden en que estén.

al conjunto de operadores que acabamos de construir. Por consiguiente, para la interacción entre un protón y un electrón en la que sólo intervienen los espines, podemos esperar que se pueda escribir el operador hamiltoniano como combinación lineal de los mismos 16 operadores. La única pregunta es, ¿cómo?

Bueno, en primer lugar, sabemos que la interacción no depende de la elección que hagamos de los ejes para un sistema de coordenadas. Si no hay perturbaciones externas —como un campo magnético— que determinen una dirección privilegiada en el espacio, el hamiltoniano no puede depender de nuestra elección de la dirección de los ejes x , y y z . Esto significa que el hamiltoniano no puede tener un término como σ_x^e completamente solo. Sería ridículo, porque entonces alguien que usara un sistema de coordenadas diferente obtendría distintos resultados.

Las únicas posibilidades son un término con la matriz unidad, digamos una constante a (por $\hat{1}$), y una combinación de las sigmas que no dependa de las coordenadas —una combinación “invariante”—. La única combinación invariante *escalar* de dos vectores es el producto escalar, el cual para nuestros σ es

$$\sigma^e \cdot \sigma^p = \sigma_x^e \sigma_x^p + \sigma_y^e \sigma_y^p + \sigma_z^e \sigma_z^p. \quad (12.4)$$

Este operador es invariante ante cualquier rotación del sistema de coordenadas. Por lo que la única posibilidad para un hamiltoniano con la simetría espacial apropiada es una constante por la matriz unidad más una constante por ese producto escalar, digamos,

$$\hat{H} = E_0 + A \sigma^e \cdot \sigma^p. \quad (12.5)$$

Éste es nuestro hamiltoniano. Es lo único que puede ser, a causa de la simetría del espacio, *mientras no haya un campo externo*. El término constante no nos dice mucho; sólo depende del nivel queelijamos como origen para medir energías. Podemos sin ningún inconveniente elegir $E_0 = 0$. El segundo término nos dice todo lo que necesitamos saber para encontrar la separación de los niveles del hidrógeno.

Si quieren, pueden pensar en el hamiltoniano de un modo diferente. Si hay dos imanes cerca uno de otro con momentos magnéticos μ_e y μ_p , la energía mutua dependerá de $\mu_e \cdot \mu_p$ —entre otras cosas—. Y, como recuerdan, encontramos que la cuestión clásica que llamamos μ_e aparece en la mecánica cuántica como $\mu_e \sigma^e$. Análogamente, lo que aparece clásicamente como μ_p aparecerá, comúnmente, en la mecánica cuántica como $\mu_p \sigma^p$ (donde μ_p es el momento magnético del protón, que es alrededor de 1000 veces más pequeño que μ_e , y tiene signo opuesto). Entonces, la ecuación (12.5) dice que la energía de interacción es como la interacción entre dos imanes —no del todo, porque la interacción entre dos imanes depende de la distancia entre ellos—. Pero la ecuación (12.5) podría ser —y, en verdad, *es*— un tipo de interacción promedio. El electrón se está moviendo por todas partes dentro del átomo, y nuestro hamiltoniano da solamente la energía promedio de interacción. Lo que dice es que para una determinada disposición espacial del electrón y el protón, hay una energía proporcional al coseno del ángulo que forman los dos momentos magnéticos, hablando clásicamente. Esta imagen cualitativa clásica les puede ayudar a comprender de dónde proviene, pero lo importante es que la ecuación (12.5) es la fórmula cuántica correcta.

El orden de magnitud de la interacción clásica entre dos imanes sería el producto de los dos momentos magnéticos dividido entre el cubo de la distancia entre ellos. La distancia entre el electrón y el protón en el átomo de hidrógeno es, hablando de manera aproximada, la mitad del radio atómico, o sea 0.5 ángstroms. De ahí que sea posible hacer la estimación burda de que la constante A es más o menos igual al producto de los dos momentos magnéticos μ_e y μ_p dividida entre el cubo de 1/2 ángstrom. Esta estimación da un número que está en el orden correcto. Se puede calcular A en forma precisa, una vez que se comprenda la teoría cuántica completa del átomo de hidrógeno —que no comprendemos aún—. De hecho, ha sido calculada con una precisión de aproximadamente 30 en un millón. Así que, contrariamente a lo que ocurría con la constante A del cambio de posición en la molécula de amoníaco, que no se podía calcular bien teóricamente,

nuestra constante A para el hidrógeno *se puede* calcular con una teoría más detallada. Pero no importa, para lo que necesitamos basta con considerar que A es un número que se podría determinar experimentalmente, y limitarnos al aspecto físico de la situación.

Con el hamiltoniano de la ecuación (12.5), podemos utilizarlo en la ecuación

$$i\hbar\dot{C}_i = \sum_j H_{ij}C_j \quad (12.6)$$

para encontrar qué es lo que hacen las interacciones de espín a los niveles de energía. Para ello, necesitamos calcular los dieciséis elementos matriciales $H_{ij} = \langle i | H | j \rangle$ correspondientes a cada par de los cuatro estados base (12.1).

Comenzaremos calculando cuánto vale $\hat{H}|j\rangle$ para cada uno de los cuatro estados base. Por ejemplo,

$$\hat{H}|++\rangle = A\sigma^e \cdot \sigma^p |++\rangle = A\{\sigma_x^e\sigma_x^p + \sigma_y^e\sigma_y^p + \sigma_z^e\sigma_z^p\}|++\rangle. \quad (12.7)$$

Empleando el método que describimos un poco antes —es fácil si han memorizado la tabla 12-1—, encontramos lo que cada par de σ hace a $|++\rangle$. La respuesta es

$$\begin{aligned} \sigma_x^e\sigma_x^p |++\rangle &= +|--\rangle, \\ \sigma_y^e\sigma_y^p |++\rangle &= -|--\rangle, \\ \sigma_z^e\sigma_z^p |++\rangle &= +|++\rangle. \end{aligned} \quad (12.8)$$

Entonces la ecuación (12.7) se convierte en

$$\hat{H}|++\rangle = A\{|--\rangle - |--\rangle + |++\rangle\} = A|++\rangle. \quad (12.9)$$

Como nuestros cuatro estados base son ortogonales, obtenemos inmediatamente:

$$\begin{aligned} \langle ++ | H | ++ \rangle &= A \langle ++ | ++ \rangle = A, \\ \langle +- | H | ++ \rangle &= A \langle +- | ++ \rangle = 0, \\ \langle -+ | H | ++ \rangle &= A \langle -+ | ++ \rangle = 0, \\ \langle -- | H | ++ \rangle &= A \langle -- | ++ \rangle = 0. \end{aligned} \quad (12.10)$$

Recordando que $\langle j | H | i \rangle = \langle i | H | j \rangle^*$, ya podemos escribir la ecuación diferencial para la amplitud C_1 :

$$i\hbar\dot{C}_1 = H_{11}C_1 + H_{12}C_2 + H_{13}C_3 + H_{14}C_4$$

o

$$i\hbar\dot{C}_1 = AC_1. \quad (12.11)$$

¡Eso es todo! Obtenemos un solo término.

Para obtener el resto de las ecuaciones hamiltonianas, tenemos que repetir una y otra vez el mismo procedimiento con $\hat{H}$ operando sobre los otros estados. Primero, los dejaremos practicar verificando todos los productos de sigmas que hemos escrito en la tabla 12-2. Luego podemos usarlos para obtener:

$$\begin{aligned} \hat{H}|+-\rangle &= A\{2|+-\rangle - |++\rangle\}, \\ \hat{H}|-\rangle &= A\{2|+-\rangle - |++\rangle\}, \\ \hat{H}|--\rangle &= A|--\rangle. \end{aligned} \quad (12.12)$$

Luego, multiplicando cada uno a la izquierda por los demás vectores de estado, obtenemos la siguiente matriz hamiltoniana, H_{ij} :

$$H_{ij} = \begin{matrix} & \begin{matrix} j \rightarrow \\ \downarrow i \end{matrix} \\ \begin{matrix} A & 0 & 0 & 0 \\ 0 & -A & 2A & 0 \\ 0 & 2A & -A & 0 \\ 0 & 0 & 0 & A \end{matrix} \end{matrix}. \quad (12.13)$$

Tabla 12-2

Operadores de espín para el átomo de hidrógeno

$\sigma_x^e\sigma_x^p ++\rangle = + --\rangle$
$\sigma_x^e\sigma_x^p +-\rangle = + -+\rangle$
$\sigma_x^e\sigma_x^p -+\rangle = + +-\rangle$
$\sigma_x^e\sigma_x^p --\rangle = + ++\rangle$
$\sigma_y^e\sigma_y^p ++\rangle = - --\rangle$
$\sigma_y^e\sigma_y^p +-\rangle = + -+\rangle$
$\sigma_y^e\sigma_y^p -+\rangle = + +-\rangle$
$\sigma_y^e\sigma_y^p --\rangle = - ++\rangle$
$\sigma_z^e\sigma_z^p ++\rangle = + ++\rangle$
$\sigma_z^e\sigma_z^p +-\rangle = - -+\rangle$
$\sigma_z^e\sigma_z^p -+\rangle = - -+\rangle$
$\sigma_z^e\sigma_z^p --\rangle = + --\rangle$

Naturalmente, esto equivale a decir que nuestras ecuaciones diferenciales para las cuatro amplitudes C_i son

$$\begin{aligned} i\hbar\dot{C}_1 &= AC_1, \\ i\hbar\dot{C}_2 &= -AC_2 + 2AC_3, \\ i\hbar\dot{C}_3 &= 2AC_2 - AC_3, \\ i\hbar\dot{C}_4 &= AC_4. \end{aligned} \tag{12.14}$$

Antes de resolver estas ecuaciones, no podemos resistir la tentación de presentarles una regla ingeniosa de Dirac —les dará la impresión de que están realmente avanzados— aunque no la necesitamos para nuestro trabajo. Tenemos —de las ecuaciones (12.9) y (12.12)— que

$$\begin{aligned} \sigma^e \cdot \sigma^p |++\rangle &= |++\rangle, \\ \sigma^e \cdot \sigma^p |+-\rangle &= 2|+-\rangle - |+-\rangle, \\ \sigma^e \cdot \sigma^p |-+\rangle &= 2|-+\rangle - |-+\rangle, \\ \sigma^e \cdot \sigma^p |--\rangle &= |--\rangle. \end{aligned} \tag{12.15}$$

Miren, dijo Dirac, también puedo escribir la primera y la última ecuación así:

$$\begin{aligned} \sigma^e \cdot \sigma^p |++\rangle &= 2|++\rangle - |++\rangle, \\ \sigma^e \cdot \sigma^p |--\rangle &= 2|--\rangle - |--\rangle; \end{aligned}$$

entonces todas son bastante similares. Ahora invento un nuevo operador, que llamaré $P_{\text{int espín}}$ y que *defino* como teniendo las siguientes propiedades:[†]

$$\begin{aligned} P_{\text{int espín}} |++\rangle &= |++\rangle, \\ P_{\text{int espín}} |+-\rangle &= |-+\rangle, \\ P_{\text{int espín}} |-+\rangle &= |+-\rangle, \\ P_{\text{int espín}} |--\rangle &= |--\rangle. \end{aligned}$$

Todo lo que el operador hace es intercambiar las direcciones del espín de las dos partículas. Luego, puedo escribir el conjunto de ecuaciones (12.15) como una simple ecuación entre operadores:

$$\sigma^e \cdot \sigma^p = 2P_{\text{int espín}} - 1. \tag{12.16}$$

Ésta es la fórmula de Dirac. Su “operador de intercambio de espín” da una regla cómoda para determinar $\sigma^e \cdot \sigma^p$. (Como ven, ya pueden hacer lo que se les ocurra. Las puertas están abiertas.)

12-3 Los niveles de energía

Ahora estamos en condiciones para encontrar los niveles de energía del estado fundamental del hidrógeno, resolviendo las ecuaciones hamiltonianas (12.14). Queremos determinar las energías de los estados estacionarios. Esto significa que queremos encontrar los estados especiales $|\psi\rangle$ para los cuales cada amplitud $C_i = \langle i | \psi \rangle$ del conjunto perteneciente a $|\psi\rangle$ tiene la misma dependencia temporal —digamos, $e^{-i\omega t}$ —. Luego, el estado tendrá la energía $E = \hbar\omega$. Por lo que queremos un conjunto para el cual

$$C_i = a_i e^{-(i/\hbar)Et}, \tag{12.17}$$

donde los cuatro coeficientes a_i son independientes del tiempo. Para ver si podemos obtener esas amplitudes, sustituimos (12.17) en la ecuación (12.14) y

[†] Actualmente, este operador se denomina “operador de intercambio de espín de Pauli”.

vemos qué pasa. Cada $i\hbar dC/dt$ en la ecuación (12.14) se convierte en EC , y —después de cancelar el factor exponencial común— cada C se vuelve una a ; obtenemos

$$\begin{aligned} Ea_1 &= Aa_1, \\ Ea_2 &= -Aa_2 + 2Aa_3, \\ Ea_3 &= 2Aa_2 - Aa_3, \\ Ea_4 &= Aa_4, \end{aligned} \tag{12.18}$$

que tenemos que resolver para a_1 , a_2 , a_3 y a_4 . No es una fortuna que la primera ecuación sea independiente del resto —esto significa que ya tenemos directamente una solución—. Si elegimos $E = A$,

$$a_1 = 1, \quad a_2 = a_3 = a_4 = 0,$$

da una solución. (Es claro que tomar todas las a iguales a cero también es una solución, ¡pero ese no es ningún estado!) Llamemos a nuestra primera solución el estado $|I\rangle$.[†]

$$|I\rangle = |1\rangle = |++\rangle. \tag{12.19}$$

Su energía es

$$E_I = A.$$

Con esa pista ustedes pueden ver inmediatamente otra solución a partir de la última ecuación en (12.18):

$$\begin{aligned} a_1 = a_2 = a_3 &= 0, & a_4 &= 1, \\ E &= A. \end{aligned}$$

Llamaremos a esta solución el estado $|II\rangle$:

$$\begin{aligned} |II\rangle &= |4\rangle = |--\rangle, \\ E_{II} &= A. \end{aligned} \tag{12.20}$$

Ahora la cuestión se torna un poco más difícil; las dos ecuaciones que quedan en (12.18) están mezcladas. Pero ya se nos ha presentado esto antes. Sumando ambas, obtenemos

$$E(a_2 + a_3) = A(a_2 + a_3). \tag{12.21}$$

Restándolas, obtenemos

$$E(a_2 - a_3) = -3A(a_2 - a_3). \tag{12.22}$$

Vemos por simple inspección —y recordando al amoníaco— que hay dos soluciones:

$$\begin{aligned} a_2 &= a_3, & E &= A \end{aligned} \tag{12.23}$$

y

$$a_2 = -a_3, \quad E = -3A.$$

Son mezclas de $|2\rangle$ y $|3\rangle$. Llamando a estos estados $|III\rangle$ y $|IV\rangle$, y agregando el factor $1/\sqrt{2}$ para normalizar adecuadamente a los estados, tenemos

$$\begin{aligned} |III\rangle &= \frac{1}{\sqrt{2}} (|2\rangle + |3\rangle) = \frac{1}{\sqrt{2}} (|+-\rangle + |-+\rangle), \\ E_{III} &= A \end{aligned} \tag{12.24}$$

[†] En realidad, el estado es $|I\rangle e^{-(i/\hbar)E_I t}$, pero como de costumbre, identificaremos a los estados mediante los vectores constantes, que son iguales a los vectores completos en $t = 0$.

y

$$|IV\rangle = \frac{1}{\sqrt{2}}(|2\rangle - |3\rangle) = \frac{1}{\sqrt{2}}(|+-\rangle - |-+\rangle), \quad (12.25)$$

$$E_{IV} = -3A.$$

Hemos encontrado cuatro estados estacionarios y sus correspondientes energías. A propósito, observen que nuestros cuatro estados son ortogonales, por lo que se pueden usar como estados base si lo deseamos. Nuestro problema está completamente resuelto.

Tres de los estados tienen energía A , y el último tiene la energía $-3A$. El promedio es cero —lo cual significa que al elegir $E_0 = 0$ en la ecuación (12.5), estábamos eligiendo medir las energías a partir de la energía promedio—. Podemos dibujar el diagrama de niveles de energía para el estado fundamental del hidrógeno como se muestra en la figura 12-2.

Ahora bien, la diferencia de energía entre el estado $|IV\rangle$ y cualquiera de los otros es $4A$. Un átomo que hubiera alcanzado el estado $|I\rangle$, podría caer al estado $|IV\rangle$ y emitir luz. No luz visible, porque la energía es muy pequeña —emitirá un cuanto de microondas—. O, si iluminamos con microondas al gas hidrógeno, observaremos absorción de energía debido a que los átomos que se encuentran en el estado $|IV\rangle$ toman energía y pasan a uno de los estados superiores —pero sólo a la frecuencia $\omega = 4A/\hbar$ —. Esta frecuencia ha sido medida experimentalmente; el mejor resultado, obtenido muy recientemente,[†] es

$$f = \omega/2\pi = (1\,420\,405\,751.800 \pm 0.028) \text{ ciclos por segundo}. \quad (12.26)$$

¡El error es de sólo dos mil millonésimas! Probablemente, ninguna cantidad física básica se ha medido mejor —es una de las mediciones más notablemente exactas de la física—. Los teóricos estaban muy contentos de poder calcular la energía con una exactitud de 3 en 10^5 , pero, mientras tanto, ha sido medida con un error de 2 en 10^{11} —un millón de veces más exacta que la teoría—. Así que los experimentales están mucho más adelantados que los teóricos. En la teoría del estado fundamental del átomo de hidrógeno, *ustedes* son tan buenos como cualquiera. También ustedes pueden sencillamente tomar el valor experimental de A —es lo que finalmente hace todo el mundo—.

Probablemente, han oído hablar de la “línea de veintidós centímetros” del hidrógeno. Ésa es la longitud de onda de la línea espectral de 1 420 megaciclos entre los estados hiperfinos. Radiación de esta longitud de onda es emitida o absorbida por el hidrógeno atómico gaseoso de las galaxias. Por ello, utilizando radiotelescopios sintonizados para ondas de 21 cm (o sea de 1 420 megaciclos aproximadamente) podemos observar las velocidades y ubicación de las concentraciones de hidrógeno atómico gaseoso. Al medir la intensidad, podemos estimar la cantidad de hidrógeno. Al medir la variación en la frecuencia debida al efecto Doppler, podemos obtener información sobre el movimiento del gas en la galaxia. Éste es uno de los grandes programas de la radioastronomía. Así que estamos hablando de algo muy real —no es un problema artificial—.

12-4 La separación de Zeeman

Aunque hemos concluido con el problema de encontrar los niveles de energía para el estado fundamental del hidrógeno, nos gustaría estudiar un poco más este sistema tan interesante. Para decir algo más sobre él —por ejemplo, para calcular la rapidez a la que el átomo de hidrógeno absorbe o emite ondas de radio de 21 cm— tenemos que saber lo que ocurre cuando se perturba al átomo. Debemos proceder como con la molécula de amoníaco —después que encontramos los niveles de energía, seguimos adelante y estudiamos lo que ocurría cuando la molécula estaba en un campo eléctrico—. Así pudimos calcular los efectos del campo eléctrico en una onda de radio. En el caso del átomo de hidrógeno, el

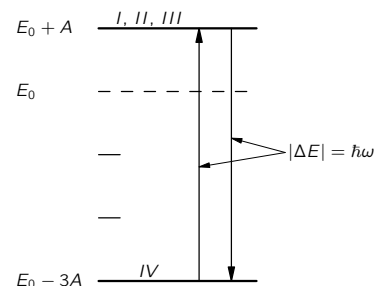


Figura 12-2. Diagrama de niveles de energía para el estado fundamental del hidrógeno atómico.

[†] Crampton, Kleppner y Ramsey, *Physical Review Letters*, vol. 11, p. 338 (1963).

campo eléctrico no hace nada a los niveles, excepto moverlos a todos una cantidad constante, proporcional al cuadrado del campo —lo cual no es interesante porque no cambia las *diferencias* de energía—. Ahora es el campo *magnético* lo que es importante. Así, el próximo paso es escribir el hamiltoniano para una situación más complicada en la que el átomo se encuentra en un campo magnético externo.

¿Cuál es, entonces, el hamiltoniano? Sólo les daremos la respuesta, porque no les podemos dar ninguna “demostración”, excepto decir que ésa es la forma en que el átomo funciona.

El hamiltoniano es

$$\hat{H} = A(\boldsymbol{\sigma}^e \cdot \boldsymbol{\sigma}^p) - \mu_e \boldsymbol{\sigma}^e \cdot \mathbf{B} - \mu_p \boldsymbol{\sigma}^p \cdot \mathbf{B}. \quad (12.27)$$

Ahora consta de tres partes. El primer término, $A\boldsymbol{\sigma}^e \cdot \boldsymbol{\sigma}^p$, representa la interacción magnética entre el electrón y el protón —es el mismo que estaría presente si no hubiera campo magnético—. Éste es el término que ya teníamos; y la influencia del campo magnético sobre la constante A es despreciable. El efecto del campo magnético externo se manifiesta a través de los dos últimos términos. El segundo término, $-\mu_e \boldsymbol{\sigma}^e \cdot \mathbf{B}$, es la energía que tendría el electrón en el campo magnético si estuviera solo.† Del mismo modo, el último término, $-\mu_p \boldsymbol{\sigma}^p \cdot \mathbf{B}$, habría sido la energía de un protón solo. Clásicamente, la energía de ambos sería la suma de las dos, y lo mismo sucede de manera cuántica. En un campo magnético, la energía de interacción debida al campo magnético es la suma de la energía de interacción del electrón con el campo externo, y la del protón con el campo —ambas expresadas en términos de los operadores sigma—. En mecánica cuántica, estos términos no son realmente las energías, pero pensar en las fórmulas clásicas de la energía es un modo para recordar las reglas para obtener el hamiltoniano. De todas maneras, la ecuación (12.27) es el hamiltoniano correcto.

Ahora tenemos que volver al principio y resolver todo el problema de nuevo. Sin embargo, ya hay mucho trabajo adelantado —sólo necesitamos agregar los efectos de los nuevos términos—. Tomemos un campo magnético constante $\mathbf{B}$ en dirección z . Luego, tenemos que sumar a nuestro operador hamiltoniano $\hat{H}$ las dos nuevas partes —las cuales podemos llamar $\hat{H}'$ —:

$$\hat{H}' = -(\mu_e \sigma_z^e + \mu_p \sigma_z^p)B.$$

Empleando la tabla 12-1, obtenemos inmediatamente:

$$\begin{aligned} \hat{H}' |++\rangle &= -(\mu_e + \mu_p)B |++\rangle, \\ \hat{H}' |+-\rangle &= -(\mu_e - \mu_p)B |+-\rangle, \\ \hat{H}' |-+\rangle &= -(-\mu_e + \mu_p)B |-+\rangle, \\ \hat{H}' |--\rangle &= (\mu_e + \mu_p)B |--\rangle. \end{aligned} \quad (12.28)$$

¡Cuán conveniente! Cuando $\hat{H}'$ opera sobre cada estado, da simplemente un número multiplicado por ese estado. La matriz $\langle i | H' | j \rangle$ tiene, por lo tanto, sólo elementos *diagonales* —podemos simplemente sumar los coeficientes que aparecen en (12.28) a los correspondientes términos diagonales en (12.13)—, y las ecuaciones hamiltonianas (12.14) se convierten en

$$\begin{aligned} i\hbar dC_1/dt &= \{A - (\mu_e + \mu_p)B\}C_1, \\ i\hbar dC_2/dt &= -\{A + (\mu_e - \mu_p)B\}C_2 + 2AC_3, \\ i\hbar dC_3/dt &= 2AC_2 - \{A - (\mu_e - \mu_p)B\}C_3, \\ i\hbar dC_4/dt &= \{A + (\mu_e + \mu_p)B\}C_4. \end{aligned} \quad (12.29)$$

† Recuerden que clásicamente $U = -\boldsymbol{\mu} \cdot \mathbf{B}$, por lo que la energía es mínima cuando el momento está en la dirección del campo. Para partículas positivas, el momento magnético es paralelo al espín, y para partículas negativas es opuesto. Por lo tanto, en la ecuación (12.27), μ_p es un número *positivo*, pero μ_e es un número *negativo*.

La forma de las ecuaciones no es diferente —sólo los coeficientes son distintos—. Por lo que, mientras B no varíe en el tiempo, podemos continuar como antes. Sustituyendo $C_i = a_i e^{-(i/\hbar)Et}$, obtenemos —como modificación de (12.18)—

$$\begin{aligned} Ea_1 &= \{A - (\mu_e + \mu_p)B\}a_1, \\ Ea_2 &= -\{A + (\mu_e - \mu_p)B\}a_2 + 2Aa_3, \\ Ea_3 &= 2Aa_2 - \{A - (\mu_e - \mu_p)B\}a_3, \\ Ea_4 &= \{A + (\mu_e + \mu_p)B\}a_4. \end{aligned} \quad (12.30)$$

Afortunadamente, la primera y cuarta ecuaciones son aún independientes del resto, por lo que de nuevo funciona la misma técnica.

Una solución es el estado $|I\rangle$, para el cual $a_1 = 1$, $a_2 = a_3 = a_4 = 0$, o sea

$$|I\rangle = |1\rangle = |++\rangle, \quad (12.31)$$

con

$$E_I = A - (\mu_e + \mu_p)B.$$

Otra es

$$|II\rangle = |4\rangle = |--\rangle, \quad (12.32)$$

con

$$E_{II} = A + (\mu_e + \mu_p)B.$$

Para las otras dos ecuaciones hay que trabajar un poco más, porque ahora los coeficientes de a_2 y a_3 ya no son iguales. Pero son justamente como el par que teníamos para la molécula de amoníaco. Recordando la ecuación (9.20), podemos hacer la siguiente analogía (teniendo en cuenta que las etiquetas 1 y 2 que aparecen ahí corresponden aquí a 2 y 3):

$$\begin{aligned} H_{11} &\rightarrow -A - (\mu_e - \mu_p)B, \\ H_{12} &\rightarrow 2A, \\ H_{21} &\rightarrow 2A, \\ H_{22} &\rightarrow -A + (\mu_e - \mu_p)B. \end{aligned} \quad (12.33)$$

Las energías están dadas entonces por (9.25), que era

$$E = \frac{H_{11} + H_{22}}{2} \pm \sqrt{\frac{(H_{11} - H_{22})^2}{4} + H_{12}H_{21}}. \quad (12.34)$$

Haciendo las sustituciones indicadas en (12.33), la fórmula para la energía es

$$E = -A \pm \sqrt{(\mu_e - \mu_p)^2 B^2 + 4A^2}.$$

Aunque en el capítulo 9 llamábamos a estas energías E_I y E_{II} , en este problema las llamamos E_{III} y E_{IV} ,

$$\begin{aligned} E_{III} &= A\{-1 + 2\sqrt{1 + (\mu_e - \mu_p)^2 B^2 / 4A^2}\}, \\ E_{IV} &= -A\{1 + 2\sqrt{1 + (\mu_e - \mu_p)^2 B^2 / 4A^2}\}. \end{aligned} \quad (12.35)$$

Hemos encontrado las energías de los cuatro estados estacionarios de un átomo de hidrógeno en un campo magnético constante. Comprobemos nuestros resultados haciendo que B tienda a cero, y viendo si obtenemos las mismas energías que teníamos en la sección anterior. Ustedes ven que realmente es así. Para $B = 0$, las energías E_I , E_{II} y E_{III} tienden a $+A$, y E_{IV} a $-3A$. Hasta los números de los estados concuerdan con los que les dimos antes. Pero cuando aplicamos el campo magnético, todas las energías cambian de manera diferente. Veamos cómo lo hacen.

En primer lugar, tenemos que recordar que para el electrón μ_e es negativo, y alrededor de 1000 veces mayor que μ_p —que es positivo—. Por lo tanto, $\mu_e + \mu_p$ y $\mu_e - \mu_p$ son números negativos y casi iguales. Llamémoslos $-\mu$ y $-\mu'$:

$$\mu = -(\mu_e + \mu_p), \quad \mu' = -(\mu_e - \mu_p). \quad (12.36)$$

(Tanto μ como μ' son números positivos, y casi iguales a la magnitud de μ_e —que es aproximadamente un magnetón de Bohr—.) Las cuatro energías son entonces:

$$\begin{aligned} E_I &= A + \mu B, \\ E_{II} &= A - \mu B, \\ E_{III} &= A\{-1 + 2\sqrt{1 + \mu'^2 B^2 / 4A^2}\}, \\ E_{IV} &= -A\{1 + 2\sqrt{1 + \mu'^2 B^2 / 4A^2}\}. \end{aligned} \quad (12.37)$$

La energía E_I comienza en A y aumenta de manera lineal con B —con pendiente μ —. La energía E_{II} también comienza en A , pero *disminuye* linealmente cuando B aumenta —su pendiente es $-\mu$ —. Estos dos niveles varían con B como muestra la figura 12-3. En esta figura aparecen también las energías E_{III} y E_{IV} . Dependen de B en forma diferente. Para B pequeña, dependen cuadráticamente de B , por lo que comienzan con pendiente horizontal. Luego empiezan a curvarse, y para B grande se aproximan a líneas rectas de pendientes $\pm\mu'$, que son casi iguales a las pendientes de E_I y E_{II} .

La variación de los niveles de energía de un átomo, debido a un campo magnético, se denomina *efecto Zeeman*. Decimos que las curvas de la figura 12-3 muestran la *separación de Zeeman* del estado fundamental del hidrógeno. Cuando no hay campo magnético, obtenemos solamente una línea espectral para la estructura hiperfina del hidrógeno. Las transiciones entre el estado $|IV\rangle$ y cualquiera de los otros se producen con absorción o emisión de un fotón cuya

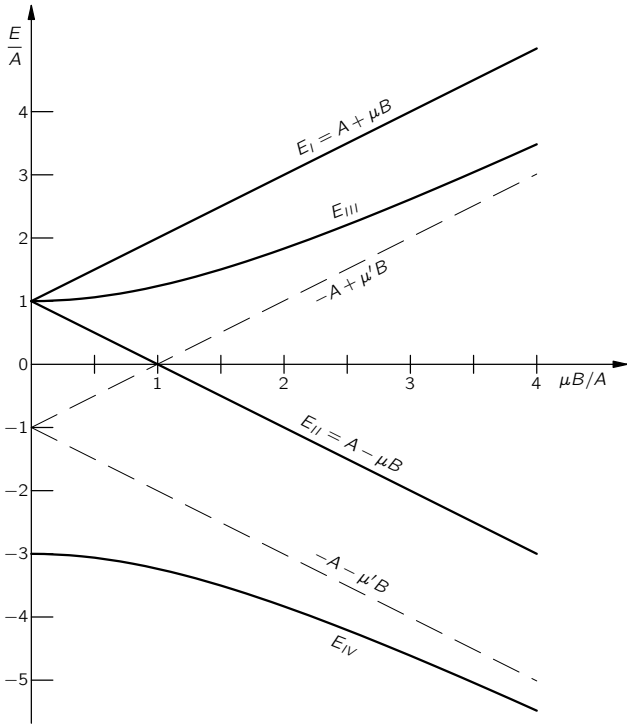


Figura 12-3. Los niveles de energía para el estado fundamental del hidrógeno en un campo magnético B .

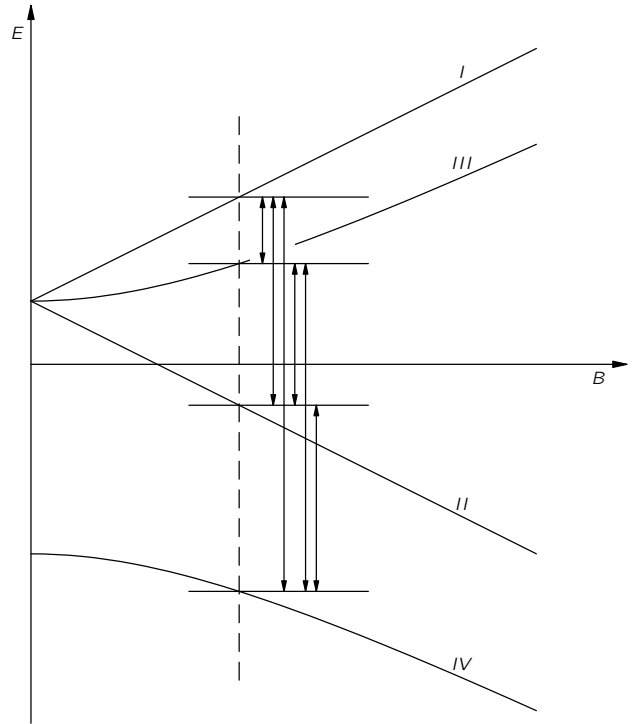


Figura 12-4. Transiciones entre los niveles de energía del estado fundamental del hidrógeno en algún campo magnético particular B .

frecuencia, 1 420 megaciclos, es $1/h$ por la diferencia de energía $4A$. Sin embargo, cuando el átomo está en un campo magnético $\mathbf{B}$, hay muchas más líneas. Puede haber transiciones entre cualesquiera dos de los cuatro estados. Así, si tenemos átomos en los cuatro estados, puede haber absorción —o emisión— de energía mediante cualesquiera de las seis transiciones indicadas con flechas verticales en la figura 12-4. Es posible observar muchas de estas transiciones con la técnica del haz molecular de Rabi que describimos en el volumen II, sección 35-3 (véase el apéndice).

¿Qué es lo que hace que haya transiciones? Las transiciones se producirán si aplican un pequeño campo magnético perturbador que varíe con el tiempo (además del intenso campo estacionario $\mathbf{B}$). Es como ocurría con un campo eléctrico variable en la molécula de amoníaco. Sólo que aquí, el campo magnético es el que se acopla con los momentos magnéticos y produce el efecto. Pero la teoría sigue los mismos lineamientos que trazamos para la molécula de amoníaco. La teoría más simple es cuando se toma un campo magnético perturbador que rota en el plano xy —aunque basta cualquier campo que oscile horizontalmente—. Al poner este campo perturbador como un término adicional en el hamiltoniano, obtienen soluciones en las que las amplitudes varían con el tiempo —como encontramos para la molécula de amoníaco—. Así entonces, ustedes pueden calcular de manera fácil y precisa la probabilidad para una transición desde un estado hasta otro. Y verán que todo concuerda con los datos experimentales.

12-5 Los estados en un campo magnético

Nos gustaría discutir ahora la forma de las curvas de la figura 12-3. Primero, las energías correspondientes a campos grandes son fáciles de comprender, y bastante interesantes. Para B suficientemente grande (es decir, para $\mu B/A \gg 1$) podemos desprestigiar el 1 en las fórmulas (12.37). Las cuatro energías se convierten en

$$\begin{aligned} E_I &= A + \mu B, & E_{II} &= A - \mu B, \\ E_{III} &= -A + \mu' B, & E_{IV} &= -A - \mu' B. \end{aligned} \quad (12.38)$$

Éstas son las ecuaciones de las cuatro rectas de la figura 12-3. Podemos entender físicamente estas energías en la siguiente forma. La naturaleza de los estados estacionarios en un campo *nulo* está completamente determinada por la interacción de los dos momentos magnéticos. Las mezclas de los estados base $|+\rangle$ y $|-\rangle$ en los estados estacionarios $|III\rangle$ y $|IV\rangle$ se deben a esta interacción. Por el contrario, en campos *externos intensos*, el protón y el electrón apenas estarán influidos por el campo del otro; cada uno se comportará como si estuviera solo en el campo externo. En consecuencia —como hemos visto muchas veces— el espín del electrón estará ya sea paralelo u opuesto al campo magnético externo.

Supongan que el espín del electrón está “hacia arriba” —es decir, en la dirección del campo—; su energía será $-\mu_e B$. El del protón puede estar de las dos maneras. Si el espín del protón también está “hacia arriba”, su energía es $-\mu_p B$. La suma de las dos es $-(\mu_e + \mu_p)B = \mu B$. Es precisamente lo que encontramos para E_I —lo cual está muy bien, porque estamos describiendo el estado $|++\rangle = |I\rangle$ —. Está aún el pequeño término adicional A (ahora $\mu B \gg A$) que representa la energía de interacción del protón y el electrón cuando sus espines son paralelos. (Al principio consideramos a A positiva porque la teoría que estábamos usando dice que así debe ser, y experimentalmente en verdad es así.) Por otra parte, el protón puede tener su espín hacia abajo. Entonces su energía en el campo externo es $+\mu_p B$, por lo que éste y el electrón tienen la energía $-(\mu_e - \mu_p)B = \mu' B$. Y la energía de interacción es $-A$. La suma es justamente la energía E_{III} en (12.38). Por lo tanto, para campos intensos, el estado $|III\rangle$ se debe transformar en el estado $|+-\rangle$.

Supongan ahora que el espín del electrón está “hacia abajo”. Su energía en el campo externo es $\mu_e B$. Si el protón también está “hacia abajo”, los dos juntos tienen la energía $(\mu_e + \mu_p)B = -\mu B$, *más* la energía de interacción A —ya que sus espines son paralelos—. Ésta es justamente la energía E_{II} en (12.38)

y corresponde al estado $| - - \rangle = | II \rangle$ —lo cual está muy bien—. Finalmente, si el electrón está “hacia abajo” y el protón “hacia arriba”, obtenemos la energía $(\mu_e - \mu_p)B - A$ (menos A para la interacción, porque los espines son opuestos), la cual es justamente E_{IV} . Y el estado corresponde a $| - + \rangle$.

“Pero, ¡espere un momento!”, estarán diciendo probablemente. “Los estados $| III \rangle$ y $| IV \rangle$ *no son* los estados $| + - \rangle$ y $| - + \rangle$; son *mezclas* de los dos.” Bueno, pero sólo en parte. Es cierto que son mezclas para $B = 0$, pero todavía no hemos visto lo que son para B grande. Al usar las analogías de (12.33) en nuestras fórmulas del capítulo 9 para obtener las energías de los estados estacionarios, también podríamos haber tomado las amplitudes correspondientes. Proviene de la ecuación (9.24), que es

$$\frac{a_2}{a_3} = \frac{E - H_{22}}{H_{21}}.$$

El cociente a_2/a_3 es, por supuesto, C_2/C_3 . Insertando las cantidades análogas indicadas en (12.33), obtenemos

$$\frac{C_2}{C_3} = \frac{E + A - (\mu_e - \mu_p)B}{2A}$$

o

$$\frac{C_2}{C_3} = \frac{E + A + \mu' B}{2A}, \quad (12.39)$$

donde, para E , tenemos que usar la energía apropiada $-E_{III}$ o E_{IV} . Por ejemplo, para el estado $| III \rangle$ tenemos

$$\left(\frac{C_2}{C_3} \right)_{III} \approx \frac{\mu' B}{A}. \quad (12.40)$$

Esto significa que, para B grande, el estado $| III \rangle$ tiene $C_2 \gg C_3$; el estado se convierte casi completamente en el estado $| 2 \rangle = | + - \rangle$. Análogamente, si ponemos E_{IV} en (12.39) obtenemos $(C_2/C_3)_{IV} \ll 1$; para campos grandes, el estado $| IV \rangle$ se convierte justamente en el estado $| 3 \rangle = | - + \rangle$. Pueden ver que los coeficientes en las combinaciones lineales de nuestros estados base que constituyen los estados estacionarios dependen de B . El estado al que llamamos $| III \rangle$ es una mezcla de $| + - \rangle$ y $| - + \rangle$ en proporciones iguales, cuando los campos son muy bajos, pero se convierte por completo en $| + - \rangle$ para campos grandes. Análogamente, el estado $| IV \rangle$, que para campos débiles es una mezcla mitad y mitad (con signos opuestos) de $| + - \rangle$ y $| - + \rangle$, se convierte en el estado $| - + \rangle$ cuando un campo externo intenso desacopla los espines.

También queremos llamarles la atención sobre lo que ocurre particularmente para campos magnéticos *muy débiles*. Hay una energía $-A$ que *no cambia* cuando aplican un campo magnético pequeño. Y hay otra energía $+A$ que se separa en tres niveles de energía diferentes cuando aplican un campo magnético pequeño. Para campos débiles, las energías varían con B como muestra la figura 12-5. Supongan que de alguna manera hemos seleccionado un montón de átomos de hidrógeno que tienen la energía $-3A$. Si los hiciéramos atravesar un aparato de Stern-Gerlach —con campos no muy fuertes— encontraríamos que pasan directamente. (Como su energía no depende de B , no hay —de acuerdo con el principio del trabajo virtual— ninguna fuerza que actúe sobre ellos en un gradiente de campo magnético.) Supongan, por el contrario, que seleccionáramos un montón de átomos con la energía $+A$, y los hiciéramos atravesar un aparato de Stern-Gerlach, digamos un aparato S . (También en este caso los campos que hay en el aparato no deben ser tan grandes como para alterar el interior del átomo, por lo cual entendemos un campo lo suficientemente pequeño como para que las energías varíen linealmente con B .) Encontraríamos *tres haces*. Los estados $| I \rangle$ y $| II \rangle$ experimentan fuerzas opuestas —sus energías varían linealmente respecto a B con pendientes $\pm\mu$ —; pero el estado $| III \rangle$ pasa libremente. Entonces estamos de vuelta en el capítulo 5. *Un átomo de hidrógeno con energía $+A$ es una partícula de espín uno*. Este estado de energía es una “partícula” para la cual

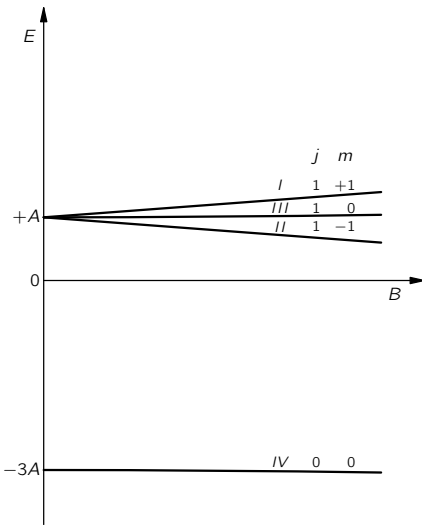


Figura 12-5. Los estados para el átomo de hidrógeno en campos magnéticos pequeños.

$j = 1$, y se puede describir —respecto a algún sistema de ejes en el espacio— en términos de los estados $|+S\rangle$, $|0S\rangle$ y $|-S\rangle$ que usamos en el capítulo 5. Por otra parte, cuando un átomo de hidrógeno tiene energía $-3A$, es una partícula de espín cero. (Recuerden que lo que estamos diciendo sólo es estrictamente cierto para campos magnéticos infinitesimales.) Por lo que podemos agrupar los estados del hidrógeno en un campo magnético nulo de esta manera:

$$\left. \begin{aligned} |I\rangle &= |++\rangle \\ |III\rangle &= \frac{|+-\rangle + |-+\rangle}{\sqrt{2}} \\ |II\rangle &= |--\rangle \end{aligned} \right\} \text{espín } 1 \quad \left\{ \begin{aligned} &|+S\rangle \\ &|0S\rangle \\ &|-S\rangle \end{aligned} \right. \quad (12.41)$$

$$|IV\rangle = \frac{|+-\rangle - |-+\rangle}{\sqrt{2}} \quad \text{espín } 0. \quad (12.42)$$

Hemos dicho en el capítulo 35 del volumen II (apéndice) que para cualquier partícula, la componente de su momentum angular en dirección de cualquier eje sólo puede tener ciertos valores, que siempre difieren en $\hbar$. La componente z del momentum angular J_z puede ser $j\hbar$, $(j-1)\hbar$, $(j-2)\hbar$, ..., $(-j)\hbar$, donde j es el espín de la partícula (que puede ser un entero o una fracción). Aunque no lo dijimos en su momento, la gente escribe comúnmente

$$J_z = m\hbar, \quad (12.43)$$

donde m es uno de los números $j, j-1, j-2, \dots, -j$. En consecuencia, verán que, en los libros, la gente designa a los cuatro estados fundamentales del hidrógeno con los llamados *números cuánticos* j y m [frecuentemente llamados “números cuánticos del momentum angular total” (j) y “número cuántico magnético” (m)]. Luego, en vez de nuestros símbolos de estado $|I\rangle$, $|II\rangle$, etc., ellos escriben un estado en la forma $|j, m\rangle$. Por lo que escribirán nuestra pequeña tabla de estados para campo nulo de las ecuaciones (12.41) y (12.42), como se muestra en la tabla 12-3. No hay ningún contenido físico nuevo, sólo es cuestión de notación.

Tabla 12-3

**Estados del átomo de hidrógeno
en un campo nulo**

Estado $ j, m\rangle$	j	m	Nuestra notación
$ 1, +1\rangle$	1	+1	$ I\rangle = +S\rangle$
$ 1, 0\rangle$	1	0	$ III\rangle = 0S\rangle$
$ 1, -1\rangle$	1	-1	$ II\rangle = -S\rangle$
$ 0, 0\rangle$	0	0	$ IV\rangle$

12-6 La matriz de proyección para espín uno†

Ahora queremos utilizar nuestro conocimiento del átomo de hidrógeno para hacer algo especial. Dijimos en el capítulo 5 que una partícula con *espín uno*, que estuviera en uno de los estados base (+, 0 o -) respecto a un aparato de Stern-Gerlach de orientación determinada —digamos un aparato S — tendría una cierta amplitud para estar en cada uno de los tres estados respecto a un aparato T con diferente orientación en el espacio. Hay nueve de esas amplitudes $\langle jT | iS \rangle$ que constituyen la matriz de proyección. En la sección 5-7 dimos sin demostración los elementos de esta matriz para diversas orientaciones de T respecto a S . Ahora mostraremos una manera de deducirlos.

En el átomo de hidrógeno encontramos un sistema de espín uno, el cual está compuesto de dos partículas con espín un medio. Ya hemos aprendido en el capítulo 6 cómo transformar las amplitudes de espín un medio. Podemos usar esta información para determinar la transformación para espín uno. Así es como se hace: tenemos un sistema —un átomo de hidrógeno con energía $+A$ — que tiene espín uno. Supongan que lo hacemos pasar por un filtro S de Stern-Gerlach, por lo que sabemos que está en uno de los estados base respecto a S , digamos en $|+S\rangle$. ¿Cuál es la amplitud para que esté en uno de los estados base, digamos $|+T\rangle$, respecto al aparato T ? Si llamamos sistema x, y, z al sistema de coordenadas del aparato S , el estado $|+S\rangle$ es lo que hemos estado llamando $|++\rangle$. Pero supongan ahora que otra persona elige su eje z en la dirección del eje de T . Estará refiriendo sus estados a lo que llamaremos sistema x', y', z' . Sus estados de espín “hacia arriba” y “hacia abajo” para el electrón y el protón serán diferentes de los nuestros. Su estado “más-más” —que podemos

† Quienes eligieron omitir el capítulo 6 también deben omitir esta sección.

escribir como $|+' +'\rangle$, referido al sistema con “primas”— es el estado $|+T\rangle$ de la partícula de espín uno. Lo que queremos es $\langle +T|+S\rangle$, que es simplemente otra manera para escribir la amplitud $\langle +' +'|++\rangle$.

Podemos encontrar la amplitud $\langle +' +'|++\rangle$ de la siguiente manera. En *nuestro* sistema de referencia, el *electrón* en el estado $|++\rangle$ tiene espín “hacia arriba”. Esto significa que tiene cierta amplitud $\langle +' |+\rangle_e$ para estar “hacia arriba” en el *otro* sistema, y cierta amplitud $\langle -' |+\rangle_e$ para estar “hacia abajo” en ese sistema. Análogamente, el *protón* en el estado $|++\rangle$ tiene espín “hacia arriba” en nuestro sistema y las amplitudes $\langle +' |+\rangle_p$ y $\langle -' |+\rangle_p$ para tener espín “hacia arriba” o “hacia abajo” en el sistema con “primas”. Como estamos hablando de dos partículas distintas, la amplitud para que *ambas* partículas estén *conjuntamente* “hacia arriba” en el sistema de *él* es el producto de las dos amplitudes,

$$\langle +' +'|++\rangle = \langle +' |+\rangle_e \langle +' |+\rangle_p. \quad (12.44)$$

Pusimos los subíndices e y p en las amplitudes $\langle +' |+\rangle$ para que quede claro lo que estamos haciendo. Pero ambas son simplemente las amplitudes de transformación para partículas de espín un medio, por lo que en realidad son números idénticos. De hecho, son justamente las amplitudes que llamamos $\langle +T|+S\rangle$ en el capítulo 6, y que dimos en las tablas que están al final de ese capítulo.

Sin embargo, estamos a punto de complicarnos con la notación. Tenemos que poder distinguir la amplitud $\langle +T|+S\rangle$ para una partícula de *espín un medio* de la que *también* llamamos $\langle +T|+S\rangle$ para una partícula de *espín uno* —¡aunque son completamente diferentes!—. Esperamos que no cause demasiada confusión, pero *por ahora* al menos, tendremos que usar símbolos diferentes para las amplitudes de espín un medio. Para ayudarlos a mantener las cosas en orden, hemos resumido la nueva notación en la tabla 12-4. Continuaremos usando la notación $|+S\rangle$, $|0S\rangle$ y $|-S\rangle$ para los estados de una partícula con espín uno.

En nuestra nueva notación, la ecuación (12.44) se convierte simplemente en

$$\langle +' +'|++\rangle = a^2,$$

y ésta es justamente la amplitud $\langle +T|+S\rangle$ para espín *uno*. Supongamos ahora, por ejemplo, que el sistema de coordenadas del otro sujeto —es decir, el aparato *T* o con “primas”— sólo está rotado respecto a *nuestro* eje *z* en un ángulo ϕ ; luego, conforme a la tabla 6-2,

$$a = \langle +' |+\rangle = e^{i\phi/2}.$$

Según (12.44) tenemos, entonces, que la amplitud para espín uno es

$$\langle +T|+S\rangle = \langle +' +'|++\rangle = (e^{i\phi/2})^2 = e^{i\phi}. \quad (12.45)$$

Ya pueden ver cómo funciona.

Ahora resolveremos el caso general para todos los estados. Si el protón y el electrón están “hacia arriba” en *nuestro* sistema —el sistema *S*— las amplitudes para que esté en *cualquiera de los cuatro* estados posibles del sistema del otro tipo —el sistema *T*— son

$$\begin{aligned} \langle +' +'|++\rangle &= \langle +' |+\rangle_e \langle +' |+\rangle_p = a^2, \\ \langle +' -'|++\rangle &= \langle +' |+\rangle_e \langle -' |+\rangle_p = ab, \\ \langle -' +'|++\rangle &= \langle -' |+\rangle_e \langle +' |+\rangle_p = ba, \\ \langle -' -'|++\rangle &= \langle -' |+\rangle_e \langle -' |+\rangle_p = b^2. \end{aligned} \quad (12.46)$$

Podemos entonces escribir el estado $|++\rangle$ como la siguiente combinación lineal:

$$|++\rangle = a^2 |+' +'\rangle + ab\{|+' -'\rangle + |-'+'\rangle\} + b^2 |-'-'\rangle. \quad (12.47)$$

Ahora notamos que $|+' +'\rangle$ es el estado $|+T\rangle$, que $|+' -'\rangle + |-'+'\rangle$ es sólo $\sqrt{2}$ por el estado $|0T\rangle$ —véase (12.41)— y que $|-'-'\rangle = |-T\rangle$. En otras palabras, podemos reescribir la ecuación (12.47) en la forma

Tabla 12-4

Amplitudes para espín un medio

Este capítulo	Capítulo 6
$a = \langle +' +\rangle$	$\langle +T +S\rangle$
$b = \langle -' +\rangle$	$\langle -T +S\rangle$
$c = \langle +' -\rangle$	$\langle +T -S\rangle$
$d = \langle -' -\rangle$	$\langle -T -S\rangle$

$$|+S\rangle = a^2 |+T\rangle + \sqrt{2} ab |0T\rangle + b^2 |-T\rangle. \quad (12.48)$$

En forma similar, pueden demostrar fácilmente que

$$|-S\rangle = c^2 |+T\rangle + \sqrt{2} cd |0T\rangle + d^2 |-T\rangle. \quad (12.49)$$

Para $|0S\rangle$ es un poco más complicado, porque

$$|0S\rangle = \frac{1}{\sqrt{2}} \{|+-\rangle + |-+\rangle\}.$$

Pero podemos expresar los estados $|+-\rangle$ y $|-+\rangle$ en términos de los estados con “primas” y sumarlos. Es decir,

$$|+-\rangle = ac |+'+\rangle + ad |+'-\rangle + bc |-' +'\rangle + bd |-' -'\rangle \quad (12.50)$$

y

$$|-+\rangle = ac |+'+\rangle + bc |+'-\rangle + ad |-' +'\rangle + bd |-' -'\rangle. \quad (12.51)$$

Multiplicando la suma por $1/\sqrt{2}$, obtenemos

$$|0S\rangle = \frac{2}{\sqrt{2}} ac |+'+\rangle + \frac{ad+bc}{\sqrt{2}} \{|+'-\rangle + |-' +'\rangle\} + \frac{2}{\sqrt{2}} bd |-' -'\rangle.$$

De aquí se deduce que

$$|0S\rangle = \sqrt{2} ac |+T\rangle + (ad+bc) |0T\rangle + \sqrt{2} bd |-T\rangle. \quad (12.52)$$

Ahora tenemos todas las amplitudes que queríamos. Los coeficientes en las ecuaciones (12.48), (12.49) y (12.52) son los elementos matriciales $\langle jT | iS \rangle$. Pongámoslos todos juntos:

$$\langle jT | iS \rangle = {}^{iT\downarrow}_{iS\rightarrow} \begin{pmatrix} a^2 & \sqrt{2} ac & c^2 \\ \sqrt{2} ab & ad+bc & \sqrt{2} cd \\ b^2 & \sqrt{2} bd & d^2 \end{pmatrix}. \quad (12.53)$$

Hemos expresado la transformación para espín uno en función de las amplitudes para espín un medio a , b , c y d .

Por ejemplo, si se rota el sistema T respecto a S en un ángulo α alrededor del eje y —como en la figura 5-6— las amplitudes de la tabla 12-4 son justamente los elementos matriciales de $R_y(\alpha)$ de la tabla 6-2.

$$\begin{aligned} a &= \cos \frac{\alpha}{2}, & b &= -\sin \frac{\alpha}{2}, \\ c &= \sin \frac{\alpha}{2}, & d &= \cos \frac{\alpha}{2}. \end{aligned} \quad (12.54)$$

Usando estas expresiones en (12.53), obtenemos las fórmulas de (5.38), que dimos entonces sin demostración.

¿Qué pasó con el estado $|IV\rangle$?! Bueno, es un sistema de espín cero, por lo que sólo tiene un estado —es el *mismo en todos los sistemas de coordenadas*—. Podemos verificar que todo está bien, restando las ecuaciones (12.50) y (12.51); obtenemos que

$$|+-\rangle - |-+\rangle = (ad-bc)\{|+'-\rangle - |-' +'\rangle\}.$$

Pero $(ad-bc)$ es el determinante de la matriz para espín un medio, por lo que es igual a 1. Obtenemos que

$$|IV'\rangle = |IV\rangle$$

para cualquier orientación relativa de los dos sistemas de coordenadas.

Propagación en una red cristalina

13-1 Estados para un electrón en una red unidimensional

Ustedes pensarían, a primera vista, que un electrón de baja energía tendría gran dificultad de atravesar un cristal sólido. Los átomos se encuentran agrupados con sus centros separados por unos pocos Angstroms solamente, y el diámetro efectivo del átomo para la dispersión del electrón es aproximadamente de un Angstrom. Esto es, los átomos son grandes en relación con su espaciamiento, de tal modo que esperarían que el camino libre medio entre colisiones fuese del orden de algunos Angstroms—que es prácticamente nada. Podrían esperar que el electrón chocara con un átomo u otro casi inmediatamente. No obstante, es un fenómeno muy frecuente en la naturaleza que, si la red es perfecta, los electrones pueden viajar a través del cristal de manera suave y fácilmente—casi como si estuviesen en el vacío. Este hecho extraño es el que permite que los metales conduzcan la electricidad tan fácilmente; también ha permitido el desarrollo de muchos dispositivos prácticos. Es, por ejemplo, lo que hace posible que un transistor imite un tubo de radio. En un tubo de radio los electrones se mueven libremente a través del vacío, mientras que en un transistor se mueven libremente a través de una red cristalina. En este capítulo describiremos la maquinaria responsable del comportamiento de un transistor; en el próximo describiremos la aplicación de estos principios a diversos dispositivos prácticos.

La conducción de electrones en un cristal es un ejemplo de un fenómeno muy común. No solamente los electrones pueden viajar a través de los cristales, sino también otras “cosas”, como las excitaciones atómicas, también pueden moverse en una forma semejante. Por eso, el fenómeno que queremos discutir aparece de muchas maneras en el estudio de la física del estado sólido.

Recordarán que hemos discutido muchos ejemplos de sistemas de dos estados. Ahora consideremos a un electrón, el cual puede estar en alguna de dos posiciones, en cada una de las cuales hay el mismo ambiente. Supongamos también que hay una cierta amplitud para ir de una posición a la otra, y, por supuesto, la misma amplitud para volver, justo como discutimos para el ion de la molécula de hidrógeno en la sección 10-1. Las leyes de la mecánica cuántica nos dan los siguientes resultados. Hay dos estados posibles con energía definida para el electrón. Cada estado se puede describir por la amplitud para que el electrón se encuentre en cada una de las dos posiciones básicas. En cualquiera de los estados de energía definida, las magnitudes de estas dos amplitudes son constantes en el tiempo, y las fases varían en el tiempo con la misma frecuencia. Por otra parte, si empezamos con el electrón en una posición, después habrá pasado a la otra, y más tarde habrá vuelto a la primera posición. La amplitud es análoga a los movimientos de dos péndulos acoplados.

Ahora consideremos una red cristalina perfecta, en la que imaginamos que podemos situar a un electrón en una especie de “pozo”, en un átomo determinado y con cierta energía determinada. Supongan también que el electrón tiene cierta amplitud para pasar a un pozo diferente, situado en alguno de los átomos cercanos. Es algo que se asemeja al sistema de dos estados—pero con una complicación adicional. Cuando el electrón llega al átomo vecino, después puede pasar a otra posición o volver a su punto de partida. Ahora tenemos una situación que no es análoga a *dos* péndulos acoplados, sino a un *número infinito* de péndulos acoplados entre sí. Es algo parecido a lo que ven en una de esas máquinas—hechas con una larga hilera de barras montadas en un hilo con cierta torsión—que se usan en el primer año de física para demostrar la propagación de las ondas.

- 13-1 Estados para un electrón en una red unidimensional**
- 13-2 Estados de energía definida**
- 13-3 Estados dependientes del tiempo**
- 13-4 Un electrón en una red tridimensional**
- 13-5 Otros estados en una red**
- 13-6 Dispersión por imperfecciones en la red**
- 13-7 Atrampar con una imperfección en la red**
- 13-8 Amplitudes de dispersión y estados enlazados**

Si tienen un oscilador armónico acoplado a otro oscilador armónico, y éste a otro y así sucesivamente, si provocan una irregularidad en un lugar, la irregularidad se propagará como una onda a lo largo de la hilera. La misma situación sucede si colocan a un electrón en un átomo de una larga cadena de átomos.

Usualmente, el modo más sencillo de analizar el problema mecánico es no pensar en términos de lo que sucede si un pulso se inicia en un lugar determinado, sino más bien en términos de soluciones de ondas estacionarias. Existen ciertos patrones de desplazamientos que se propagan a través del cristal como una onda de frecuencia fija y única. Lo mismo sucede con el electrón —y por la misma razón, porque en la mecánica cuántica se describe mediante ecuaciones parecidas—.

Sin embargo, deben apreciar una cosa: La amplitud para que el electrón esté en un lugar es una *amplitud*, no una probabilidad. Si los electrones estuviesen simplemente escurriéndose de un lugar a otro, como agua atravesando un agujero, el comportamiento sería completamente diferente. Por ejemplo, si tuviésemos dos tanques de agua unidos mediante un tubo para permitir que el agua escurriese de uno a otro, entonces los niveles se aproximarían uno a otro exponencialmente. Pero para los electrones, lo que sucede es un escurrimiento de amplitud, y no un simple escurrimiento de probabilidad. Y esto es una característica del término imaginario—la i en las ecuaciones diferenciales de la mecánica cuántica—que cambia la solución exponencial a una solución oscilatoria. Lo que sucede entonces es bastante diferente del escurrimiento entre tanques conectados.

Ahora queremos analizar cuantitativamente la situación cuántica. Imaginen un sistema unidimensional, formado por una larga línea de átomos como muestra la figura 13-1(a). (Naturalmente, un cristal es tridimensional, pero la física es muy parecida; una vez que hayan entendido el caso unidimensional, serán capaces de entender lo que sucede en tres dimensiones.) Luego, queremos ver qué sucede si colocamos un solo electrón en esta línea de átomos. Por supuesto, en un cristal real ya hay millones de electrones. Pero la mayoría de ellos (casi todos en un cristal aislado) escogen posiciones en algún patrón de movimiento, cada uno alrededor de su átomo—y el conjunto es bastante estacionario. Sin embargo, nosotros queremos saber qué sucede si colocamos un electrón *adicional*. No consideraremos lo que los otros están haciendo, porque suponemos que cambiar sus movimientos implica una gran cantidad de energía de excitación. Vamos a añadir un electrón como para producir un ion negativo levemente enlazado. Al observar lo que el *único* electrón adicional hace, vamos a hacer una aproximación, la cual no considera la mecánica del funcionamiento interno de los átomos.

Por supuesto, el electrón podría pasar a otro átomo, transfiriendo el ion negativo a otro lugar. Supondremos esto, exactamente como en el caso de un electrón saltando entre dos protones, el electrón puede saltar de un átomo al de junto en cualquier lado con una cierta amplitud.

¿Pero cómo describimos tal sistema? ¿Cuáles serán los estados base razonables? Si recuerdan lo que hicimos cuando teníamos solamente dos posiciones posibles, pueden imaginar cómo será esto. Supongan que los espaciamentos en nuestra línea de átomos son todos iguales; y que numeramos los átomos sucesivos, como muestra la figura 13-1(a). Uno de los estados base es que el electrón se encuentre en el átomo número 6, otro es que se encuentre en el átomo número 7, o en átomo número 8, y así sucesivamente. Podemos describir el n -ésimo estado base diciendo que el electrón se encuentra en el átomo número n . Digamos que éste es el estado base $|n\rangle$. La figura 13-1 muestra lo que entendemos por los tres estados base

$$|n-1\rangle, |n\rangle \text{ y } |n+1\rangle.$$

Usando estos estados base, cualquier estado $|\phi\rangle$ del electrón en nuestro cristal unidimensional se puede describir dando todas las amplitudes $\langle n|\phi\rangle$ para que el estado $|\phi\rangle$ se encuentre en uno de los estados base—lo que significa, la amplitud para que esté localizado en un átomo particular. Entonces podemos escribir al estado $|\phi\rangle$ como una superposición de los estados base

$$|\phi\rangle = \sum_n |n\rangle \langle n|\phi\rangle. \quad (13.1)$$

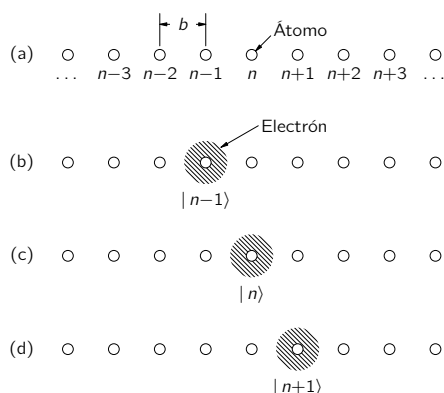


Figura 13-1. Los estados base de un electrón en un cristal unidimensional.

Después, vamos a suponer que, cuando el electrón se encuentra en un átomo, hay una cierta amplitud para que se escurra a uno de los átomos laterales. Y tomaremos el caso más simple, en que solamente se puede escurrir a los vecinos más próximos—para llegar al siguiente vecino, tiene que hacerlo en dos etapas. Diremos que las amplitudes para que el electrón salte de un átomo a otro es $iA/\hbar$ (por unidad de tiempo).

Por el momento, nos gustaría escribir la amplitud $\langle n | \phi \rangle$ para estar en el n -ésimo átomo como C_n . Así la ecuación (13.1) se escribirá

$$|\phi\rangle = \sum_n |n\rangle C_n. \quad (13.2)$$

Si conociésemos todas las amplitudes C_n en un momento dado, podríamos tomar los cuadrados de sus normas y obtener la probabilidad para encontrar al electrón si mirasen al átomo n en dicho momento.

¿Cuál será la situación después? Por analogía con los sistemas de dos estados que estudiamos, proponemos que las ecuaciones hamiltonianas para este sistema se deben construir mediante ecuaciones como ésta:

$$i\hbar \frac{dC_n(t)}{dt} = E_0 C_n(t) - AC_{n+1}(t) - AC_{n-1}(t). \quad (13.3)$$

El primer coeficiente de la derecha, E_0 , físicamente es la energía que un electrón tendría si no pudiese separarse de uno de los átomos. (No importa a qué llamemos E_0 ; como hemos visto muchas veces, no representa realmente nada, sólo nuestra elección del cero de energía.) El término siguiente representa la amplitud por unidad de tiempo para que el electrón se escurra desde el pozo n -ésimo hasta el pozo $(n+1)$ -ésimo; y el último término es la amplitud para el escurrimiento a partir del pozo $(n-1)$ -ésimo. Como de costumbre, supondremos que A es una constante (independiente de t).

Para una descripción más completa del comportamiento de cualquier estado $|\phi\rangle$, tendríamos una ecuación como (13.3) para cada una de las amplitudes C_n . Como queremos considerar un cristal con un gran número de átomos, supondremos que hay un número indefinidamente grande de estados—que los átomos no tienen fin en ambas direcciones. (Para hacer el caso finito, tendremos que prestar especial atención a lo que sucede en los átomos extremos.) Si el número N de nuestros estados base es indefinidamente grande, ¡también las ecuaciones hamiltonianas son infinitas en cantidad! Escribiremos sólo una muestra:

$$\begin{array}{ccc} \vdots & & \vdots \\ i\hbar \frac{dC_{n-1}}{dt} & = & E_0 C_{n-1} - AC_{n-2} - AC_n, \\ i\hbar \frac{dC_n}{dt} & = & E_0 C_n - AC_{n-1} - AC_{n+1}, \\ i\hbar \frac{dC_{n+1}}{dt} & = & E_0 C_{n+1} - AC_n - AC_{n+2}, \\ \vdots & & \vdots \end{array} \quad (13.4)$$

13-2 Estados de energía definida

Podríamos estudiar muchas cosas sobre un electrón en una red, pero primero tratemos de encontrar los estados de energía definida. Como hemos visto en capítulos anteriores, esto significa que tenemos que encontrar una situación en la cual todas las amplitudes cambian con la misma frecuencia, si es que varían en el tiempo. Buscamos soluciones de la forma

$$C_n = a_n e^{-iEt/\hbar}. \quad (13.5)$$

El número complejo a_n nos dice sobre la parte que no varía con el tiempo de la amplitud para encontrar al electrón en el átomo n -ésimo. Si colocamos esta

solución de prueba en las ecuaciones (13.4) para comprobarlas, obtenemos el siguiente resultado

$$Ea_n = E_0a_n - Aa_{n+1} - Aa_{n-1}. \quad (13.6)$$

Tenemos un número infinito de ecuaciones semejantes, correspondientes al número infinito de incógnitas a_n —lo cual es bastante terrorífico.

Todo lo que tenemos que hacer es calcular el determinante... ¡pero esperen! Los determinantes son estupendos cuando hay 2, 3 o 4 ecuaciones. Pero si hay un gran número—o un número infinito—de ecuaciones, los determinantes no son muy convenientes. Mejor sólo tratamos de resolver las ecuaciones directamente. Primero, numeremos a los átomos según sus *posiciones*; diremos que el átomo n se encuentra en x_n y que el átomo $(n+1)$ se encuentra en x_{n+1} . Si el espaciamiento atómico es b —como en la figura 13-1—tendremos que $x_{n+1} = x_n + b$. Incluso podríamos hacer que $x_n = nb$, escogiendo nuestro origen en el átomo cero. Podemos volver a escribir la ecuación (13.5) en la forma

$$C_n = a(x_n)e^{-iEt/\hbar}, \quad (13.7)$$

y la ecuación (13.6) se vuelve

$$Ea(x_n) = E_0a(x_n) - Aa(x_{n+1}) - Aa(x_{n-1}). \quad (13.8)$$

O, usando que $x_{n+1} = x_n + b$, podríamos también escribir

$$Ea(x_n) = E_0a(x_n) - Aa(x_n + b) - Aa(x_n - b). \quad (13.9)$$

Esta ecuación es algo similar a una ecuación diferencial. Nos dice que una cantidad, $a(x)$, en un punto, (x_n) , se relaciona con la misma cantidad física en algunos puntos vecinos, $(x_n \pm b)$. (Una ecuación diferencial relaciona el valor de una función en un punto con los valores en puntos infinitamente cercanos.) Puede ser que los métodos que comúnmente usamos para resolver ecuaciones diferenciales también sirvan aquí; intentemos.

Las ecuaciones diferenciales lineales con coeficientes constantes siempre se pueden resolver en términos de funciones exponenciales. Podemos probar la misma idea aquí; usemos como solución de prueba

$$a(x_n) = e^{ikx_n}. \quad (13.10)$$

La ecuación (13.9) queda así:

$$Ee^{ikx_n} = E_0e^{ikx_n} - Ae^{ik(x_n+b)} - Ae^{ik(x_n-b)}. \quad (13.11)$$

Podemos dividir ahora por el factor común e^{ikx_n} ; así obtenemos

$$E = E_0 - Ae^{ikb} - Ae^{-ikb}. \quad (13.12)$$

Los dos últimos términos son exactamente iguales a $(2A \cos kb)$, por lo que

$$E = E_0 - 2A \cos kb. \quad (13.13)$$

Hemos encontrado que, para *cualquier* elección de la constante k , hay una solución, cuya energía está dada por esta ecuación. Hay varias energías posibles que dependen de k , y cada k corresponde a una solución diferente. Hay un número infinito de soluciones—lo que no debe sorprendernos, puesto que comenzamos con un número infinito de estados base.

Veamos lo que significan estas soluciones. Para cada k , las a están dadas por la ecuación (13.10). Las amplitudes C_n están dadas por

$$C_n = e^{ikx_n}e^{-(i/\hbar)Et}, \quad (13.14)$$

donde deben tener presente que la energía E también depende de k , de acuerdo con la ecuación (13.13). La *dependencia espacial* de las amplitudes es e^{ikx_n} . Las amplitudes oscilan cuando pasamos de un átomo al próximo.

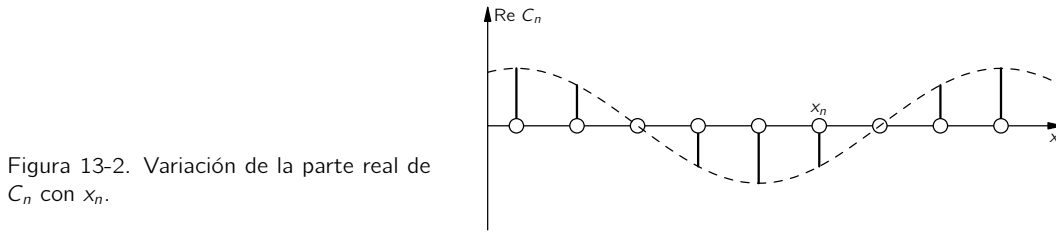


Figura 13-2. Variación de la parte real de C_n con x_n .

Queremos decir que, en el espacio, la amplitud se comporta como una oscilación *compleja*—la magnitud es la misma en cada átomo, pero la fase en un momento dado avanza en la cantidad (ikb) de un átomo al otro. Podemos visualizar lo que ocurre graficando una línea vertical que indique sólo la parte real en cada átomo, como hemos hecho en la figura 13-2. La envolvente de estas líneas verticales (como muestra la línea punteada) es, naturalmente, una curva coseno. La parte imaginaria de C_n también es una función oscilatoria, pero desfasada 90° , así que el cuadrado de la norma (que es la suma de los cuadrados de las partes real e imaginaria) es el mismo para todas las C .

Así entonces, si escogemos una k , obtenemos un estado estacionario con una energía determinada E . Y, para cualquier estado de esta clase, es igualmente probable encontrar al electrón en cada átomo—no hay preferencia por ningún átomo. Solamente la fase es diferente para átomos diferentes. También, a medida que el tiempo pasa, las fases varían. De acuerdo con la ecuación (13.14), las partes real e imaginaria se propagan en el cristal como ondas—precisamente como las partes real o imaginaria de

$$e^{i[kx_n - (E/\hbar)t]}. \quad (13.15)$$

La onda puede dirigirse hacia la x positiva o negativa, dependiendo del signo que hemos elegido para k .

Noten que hemos estado suponiendo que el número k que usamos en la solución de prueba, ecuación (13.10), es un número real. Ahora podemos ver por qué debe ser así cuando tenemos una línea infinita de átomos. Supongan que k fuese un número imaginario, digamos ik' . Entonces las amplitudes a_n serían como $e^{k'x_n}$, lo que significa que la amplitud se haría más y más grande a medida que nos dirigimos hacia x grandes—o hacia x negativas grandes, si k' es un número negativo. Esta clase de solución sería buena si estuviésemos tratando con una línea de átomos finita, pero no puede ser una solución física para una cadena infinita de átomos. Daría amplitudes infinitas—y, por lo tanto, probabilidades infinitas—que no pueden representar una situación real. Más adelante veremos un ejemplo en el que una k imaginaria tiene sentido.

La relación entre la energía E y el número de onda k mostrada en la ecuación (13.13) está graficada en la figura 13-3. Como pueden ver en la figura, la energía puede ir desde $(E_0 - 2A)$ para $k = 0$, hasta $(E_0 + 2A)$ para $k = \pm\pi/b$. La gráfica se ha hecho para A positiva; si A fuera negativa, la curva sólo estaría invertida, aunque el intervalo sería el mismo. El resultado importante es que cualquier energía es posible dentro de un cierto intervalo o “banda” de energías, pero no ninguna otra. De acuerdo con nuestras suposiciones, si un electrón en un cristal está en un estado estacionario, no puede tener otras energías que las de esta banda.

Según la ecuación (13.13), los valores de k más pequeños corresponden a estados de baja energía— $E \approx (E_0 - 2A)$. Cuando k aumenta en magnitud (hacia valores positivos o negativos) la energía crece al principio, y luego alcanza un máximo para $k = \pm\pi/b$, como muestra la figura 13-3. Para k más grandes que π/b , la energía empezaría a disminuir de nuevo. Pero realmente no necesitamos considerar esos valores de k , ya que no dan nuevos estados—sólo repiten los estados que ya teníamos para k menores. Podemos ver esto de la siguiente forma. Consideren al estado de más baja energía, para el cual $k = 0$. El coeficiente $a(x_n)$

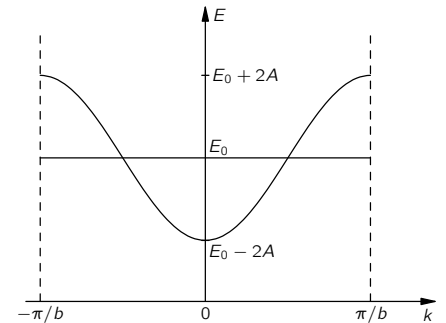


Figura 13-3. La energía de los estados estacionarios en función del parámetro k .

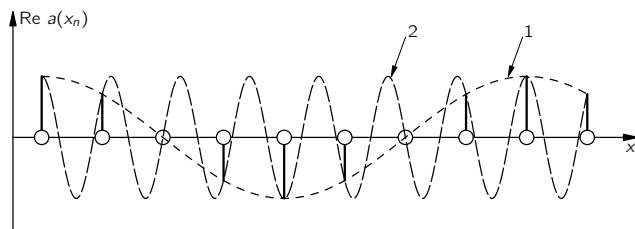


Figura 13-4. Dos valores de k que representan la misma situación física; la curva 1 es para $k = -\pi/4b$, la curva 2 es para $k = 7\pi/4b$.

es el mismo para toda x_n . Por lo tanto, obtendríamos la misma energía para $k = 2\pi/b$. Pero entonces, usando la ecuación (13.10), tenemos que

$$a(x_n) = e^{i(2\pi/b)x_n}.$$

Sin embargo, eligiendo como origen a x_0 , podemos decir que $x_n = nb$; por lo que las $a(x_n)$ se vuelven

$$a(x_n) = e^{i2\pi n} = 1.$$

El estado descrito por estas $a(x_n)$ es físicamente el mismo que teníamos para $k = 0$. No representa una solución diferente.

Como otro ejemplo, supongan que k fuese $-\pi/4b$. La parte real de $a(x_n)$ variaría como muestra la curva 1 de la figura 13-4. Si k fuese $7\pi/4b$, la parte real de $a(x_n)$ variaría como muestra la curva 2 de la figura. (Las curvas coseno completas no significan nada, naturalmente; todo lo que importa son sus valores en los puntos x_n . Las curvas sirven sólo para ayudarles a ver las cosas.) Ven que ambos valores de k dan las mismas amplitudes para todas las x_n .

Como resultado final, tenemos todas las soluciones posibles de nuestro problema si tomamos las k solamente dentro de un cierto intervalo limitado. Tomaremos el intervalo entre $-\pi/b$ y $+\pi/b$ —el que muestra la figura 13-3. En este intervalo, la energía de los estados estacionarios aumenta uniformemente con el aumento en la magnitud de k .

Una nota menor sobre algo que puede entretenerlos. Supongan que el electrón no solamente puede saltar al vecino más cercano con amplitud $iA/\hbar$, sino que también tiene la posibilidad de pasar en un salto directo al *segundo vecino* más cercano con alguna otra amplitud $iB/\hbar$. Encontrarán que la solución se puede escribir otra vez en la forma $a_n = e^{ikx_n}$ —este tipo de solución es universal. También encontrarán que los estados estacionarios con número de onda k tienen una energía igual a $(E_0 - 2A \cos kb - 2B \cos 2kb)$. Esto nos muestra que la forma de la curva para E en función de k no es universal, sino que depende de las suposiciones particulares del problema. No es siempre una onda coseno—ni siquiera es necesariamente simétrica alrededor de alguna línea horizontal. Es verdad, sin embargo, que la curva siempre se repite fuera del intervalo de $-\pi/b$ a $+\pi/b$, por lo que no necesitan preocuparse por los demás valores de k .

Consideremos un poco más a detalle lo que sucede para k pequeños—es decir, cuando las variaciones en las amplitudes de una x_n a la siguiente son bastante lentas. Supongan que escogemos nuestro cero de energía definiendo $E_0 = 2A$; así el mínimo de la curva de la figura 13-3 está en el cero de energía. Para k suficientemente pequeñas, podemos escribir

$$\cos kb \approx 1 - k^2 b^2 / 2,$$

y la energía de la ecuación (13.13) se vuelve

$$E = Ak^2 b^2. \quad (13.16)$$

Tenemos que la energía del estado es proporcional al cuadrado del número de onda que describe las variaciones espaciales de las amplitudes C_n .

13-3 Estados dependientes del tiempo

En esta sección queremos discutir el comportamiento de los estados en la red unidimensional con más detalle. Si la amplitud para que un electrón esté en x_n es C_n , la probabilidad para encontrarlo ahí es $|C_n|^2$. Para los estados *estacionarios* descritos por la ecuación (13.14), esta probabilidad es la misma para cualquier x_n y no cambia con el tiempo. ¿Cómo podemos representar una situación que describamos vagamente, diciendo que un electrón con una cierta energía está localizado en una cierta región—de tal modo que es más probable encontrarlo en un lugar que en cualquier otro? Podemos hacerlo con una superposición de varias ecuaciones semejantes a la ecuación (13.14) con valores ligeramente diferentes para k —y, por lo tanto, energías levemente distintas—. Así, por lo menos para $t = 0$, la amplitud C_n variará con la posición a causa de la interferencia entre los diversos términos, del mismo modo que se obtienen pulsaciones cuando hay una mezcla de ondas de diferentes longitudes de onda (como discutimos en el capítulo 48 del volumen I). De esta manera, podemos construir un “paquete de ondas” con un número de onda predominante k_0 , pero con otros varios números de onda cercanos a k_0 .‡

En nuestra superposición de estados estacionarios, las amplitudes con k diferentes representarán estados de energías levemente diferentes, y, por lo tanto, de frecuencias ligeramente distintas; el patrón de interferencia de la C_n total también variará, por consiguiente, con el tiempo—habrá un patrón de “pulsaciones”. Como vimos en el capítulo 48 del volumen I, los picos de las pulsaciones [los lugares donde $|C(x_n)|^2$ es mayor] se moverán a lo largo de x cuando transcurra el tiempo; se moverán con la velocidad que hemos llamado “velocidad de grupo”. Encontramos que esta velocidad de grupo está relacionada con la variación de k con la frecuencia mediante

$$v_{\text{grupo}} = \frac{d\omega}{dk}; \quad (13.17)$$

el mismo procedimiento es adecuado aquí también. Un estado del electrón que sea una “aglutinación”—es decir, un estado en que las C_n varían en el espacio como el paquete de ondas de la figura 13-5—se moverá a lo largo de nuestro “cristal” unidimensional con velocidad v igual a $d\omega/dk$, donde $\omega = E/\hbar$. Usando (13.16) para E , obtenemos

$$v = \frac{2Ab^2}{\hbar} k. \quad (13.18)$$

En otras palabras, los electrones se desplazan con una velocidad proporcional al k típico. La ecuación (13.16) indica, entonces, que la energía de tal electrón es proporcional al cuadrado de su velocidad—*actúa como una partícula clásica*. De manera que, mientras consideremos las cosas en una escala lo bastante burda para que no veamos la estructura fina, nuestra imagen cuántica empieza a darnos resultados parecidos a los de la física clásica. De hecho, si despejamos k de la ecuación (13.18) y la sustituimos en (13.16), podemos escribir

$$E = \frac{1}{2} m_{\text{ef}} v^2, \quad (13.19)$$

donde m_{ef} es una constante. La “energía de movimiento” adicional del electrón en un paquete depende de la velocidad exactamente igual que para una partícula clásica. La constante m_{ef} —llamada “masa efectiva”—está dada por

$$m_{\text{ef}} = \frac{\hbar^2}{2Ab^2}. \quad (13.20)$$

Noten también que podemos escribir

$$m_{\text{ef}} v = \hbar k. \quad (13.21)$$

Si decimos que $m_{\text{ef}} v$ es el “momentum”, éste está relacionado con el número de onda k del modo que describimos antes para una partícula libre.

No olviden que $m_{\text{ef}} v$ no tiene nada que ver con la masa real de un electrón. Puede que sea bastante diferente—aunque en los cristales reales ocurre con

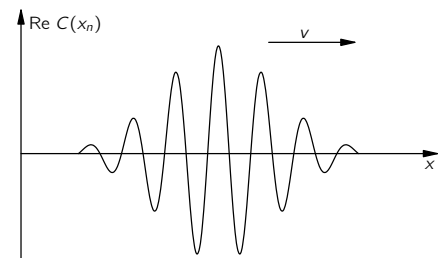


Figura 13-5. La parte real de $C(x_n)$ en función de x para una superposición de varios estados de energía similar. (El espaciamiento b es muy pequeño en la escala de x mostrada.)

‡ Siempre que no tratemos de construir el paquete demasiado estrecho.

frecuencia que es del mismo orden de magnitud, alrededor de dos a 20 veces la masa del electrón en el espacio libre.

Acabamos de explicar un misterio notable—cómo puede un electrón en un cristal (como un electrón adicional colocado en germanio) moverse directamente a través del cristal y fluir con toda libertad, aunque tenga que chocar con todos los átomos. Lo hace porque tiene sus amplitudes que van pip- pip- pip de un átomo al siguiente, abriéndose camino a través del cristal. Así es como un sólido puede conducir electricidad.

13-4 Un electrón en una red tridimensional

Consideremos por un momento cómo podríamos aplicar las mismas ideas para ver lo que sucede a un electrón en tres dimensiones. Los resultados son muy similares. Supongan que tenemos una red rectangular de átomos con espaciamientos a , b y c en las tres direcciones. (Si quieren una red cúbica, tomen los tres espaciamientos iguales.) Supongan también que la amplitud para saltar en la dirección x a un átomo vecino es $(iA_x/\hbar)$, para saltar en la dirección y es $(iA_y/\hbar)$ y para saltar en la dirección z es $(iA_z/\hbar)$ ¿cómo debemos describir ahora los estados base? Como en el caso unidimensional, un estado base es que el electrón está en el átomo cuyas coordenadas son x , y , z , donde (x, y, z) es un punto de la red. Eligiendo nuestro origen en un átomo, estos puntos se encuentran todos en

$$x = n_x a, \quad y = n_y b \quad y \quad z = n_z c,$$

donde n_x , n_y y n_z son tres números enteros cualesquiera. En lugar de usar subíndices para indicar esos puntos, ahora sólo usaremos x , y y z , entendiendo que sólo toman los valores correspondientes a puntos de la red. Así, el estado base se representa por el símbolo |electrón en x, y, z ⟩, y la amplitud para que un electrón en algún estado $|\psi\rangle$ se encuentre en este estado base es $C(x, y, z) = \langle \text{electrón en } x, y, z | \psi \rangle$.

Como antes, las amplitudes $C(x, y, z)$ pueden variar con el tiempo. Con nuestras suposiciones, las ecuaciones hamiltonianas serían así:

$$\begin{aligned} i\hbar \frac{dC(x, y, z)}{dt} = & E_0 C(x, y, z) - A_x C(x + a, y, z) - A_x C(x - a, y, z) \\ & - A_y C(x, y + b, z) - A_y C(x, y - b, z) \\ & - A_z C(x, y, z + c) - A_z C(x, y, z - c). \end{aligned} \quad (13.22)$$

Parece bastante larga, pero ustedes pueden ver de dónde proviene cada término.

Otra vez podemos tratar de encontrar un estado estacionario, en el que todas las C varíen con el tiempo de la misma manera. De nuevo, la solución es un exponencial:

$$C(x, y, z) = e^{-iEt/\hbar} e^{i(k_x x + k_y y + k_z z)}. \quad (13.23)$$

Si sustituyen esto en (13.22) ven que sirve, siempre que la energía E esté relacionada con k_x , k_y y k_z del modo siguiente:

$$E = E_0 - 2A_x \cos k_x a - 2A_y \cos k_y b - 2A_z \cos k_z c. \quad (13.24)$$

La energía depende ahora de los *tres* números de onda k_x , k_y , k_z , los cuales además son las componentes de un vector tridimensional $\mathbf{k}$. De hecho, podemos escribir la ecuación (13.23) en notación vectorial de la forma

$$C(x, y, z) = e^{-iEt/\hbar} e^{i\mathbf{k} \cdot \mathbf{r}}. \quad (13.25)$$

La amplitud varía como una *onda plana* compleja en tres dimensiones, moviéndose en la dirección de $\mathbf{k}$ y con el número de onda $k = (k_x^2 + k_y^2 + k_z^2)^{1/2}$.

La energía asociada con estos estados estacionarios depende de las tres componentes de $\mathbf{k}$ en la forma complicada dada por la ecuación (13.24). La naturaleza de la variación de E con $\mathbf{k}$ depende de los signos relativos y de las

magnitudes de A_x , A_y y A_z . Si estos tres números son positivos, y si sólo nos preocupamos por valores pequeños de k , la dependencia es relativamente simple.

Expandiendo los cosenos como hicimos para obtener la ecuación (13.16), podemos obtener ahora

$$E = E_{\text{mín}} + A_x a^2 k_x^2 + A_y b^2 k_y^2 + A_z c^2 k_z^2. \quad (13.26)$$

Para una red cúbica simple de espaciamiento a , esperaríamos que A_x , A_y y A_z fuesen iguales—por ejemplo, todas iguales a A —y tendríamos solamente

$$E = E_{\text{mín}} + Aa^2(k_x^2 + k_y^2 + k_z^2),$$

o

$$E = E_{\text{mín}} + Aa^2 k^2. \quad (13.27)$$

Ésta es justamente como la ecuación (13.16). Siguiendo el procedimiento que usamos aquí, concluiremos que un paquete de electrones en *tres* dimensiones (construido por superposición de muchos estados de energías casi iguales) también se mueve como una partícula clásica con cierta masa efectiva.

En un cristal con una simetría más baja que la cúbica (o incluso en un cristal cúbico en el que el estado del electrón en cada átomo no es simétrico) los tres coeficientes A_x , A_y y A_z son diferentes. Por lo que la “masa efectiva” de un electrón localizado en una pequeña región *depende de su dirección de movimiento*. Por ejemplo, podría tener una inercia distinta para el movimiento en la dirección x que para el movimiento en la dirección y . (Los detalles de esta situación se describen algunas veces en términos de un “tensor de masa efectiva”).

13-5 Otros estados en una red

De acuerdo con la ecuación (13.24), los estados del electrón que hemos considerado solamente pueden tener energías en una cierta “banda” de energías, que cubre el intervalo desde la energía mínima

$$E_0 - 2(A_x + A_y + A_z)$$

hasta la energía máxima

$$E_0 + 2(A_x + A_y + A_z).$$

Son posibles otras energías, pero pertenecen a una clase diferente de estados del electrón. Para los estados que hemos descrito, imaginamos estados base en los cuales se coloca un electrón en un átomo del cristal en algún estado particular, digamos el estado de energía más baja.

Si tienen un átomo en el espacio vacío, y le añaden un electrón para convertirlo en un ion, este ion se puede formar de muchas maneras. El electrón puede llegar de manera que forme el estado de energía más baja, o que forme uno u otro de los muchos “estados excitados” posibles del ion, cada uno de ellos con una energía definida por encima de la energía más baja. En un cristal puede ocurrir lo mismo. Supongan que la energía E_0 que escogimos antes corresponde a estados base que son iones con energía lo más baja posible. También podríamos imaginar un nuevo conjunto de estados base, en los cuales el electrón se coloca cerca del átomo n -ésimo de forma diferente —en uno de los estados excitados del ion— de modo que ahora la energía E_0 es un poco más alta. Como antes, hay cierta amplitud A (diferente de la anterior) para que el electrón salte de su estado excitado en un átomo al mismo estado excitado en un átomo vecino. El análisis completo se hace como anteriormente; encontramos una banda de energías posibles centradas en una energía más alta. Puede haber, en general, muchas de estas bandas, cada una corresponde a diferentes niveles de excitación.

También hay otras posibilidades. Puede haber cierta amplitud para que un electrón salte desde un estado excitado en un átomo hasta un estado desexcitado en el otro. (Esto se llama una interacción entre bandas.) La teoría matemática se hace más y más complicada cuando toman en cuenta más y más bandas, y

añaden más y más coeficientes de escurrimiento entre los estados posibles. Sin embargo, esto no incluye ninguna idea nueva; las ecuaciones se plantean de modo parecido a como hicimos en nuestro ejemplo simple.

Debemos señalar también que no hay nada más que decir acerca de los diversos coeficientes, como la amplitud A , que aparecen en la teoría. Por lo general son muy difíciles de calcular, de esta manera, en los casos de utilidad práctica, se conoce muy poco teóricamente sobre estos parámetros, y para cualquier situación real en particular sólo podemos tomar valores determinados experimentalmente.

Hay otros casos en que la física y las matemáticas son casi exactamente las que hemos encontrado para un electrón que se mueve en un cristal, pero en los cuales el “objeto” que se mueve es bastante diferente. Por ejemplo, supongan que nuestro cristal original—o mejor red lineal—fuera una línea de átomos neutros, cada uno con un electrón externo enlazado débilmente. Después imaginen que removemos un electrón. ¿Qué átomo ha perdido su electrón? Representemos ahora con C_n a la amplitud para que el electrón *esté ausente* del átomo en la posición x_n . Habrá, en general, una cierta amplitud $iA/\hbar$ para que el electrón en un átomo vecino—digamos el átomo $(n-1)$ -ésimo—salte al n -ésimo dejando al átomo $(n-1)$ -ésimo sin un electrón. Esto es lo mismo que decir que hay una amplitud A para el “electrón que falta” salte del átomo n -ésimo al átomo $(n-1)$ -ésimo. Pueden ver que las ecuaciones serán exactamente iguales—naturalmente, el valor de A no necesita ser el mismo que teníamos antes. De nuevo obtendremos las mismas fórmulas para los niveles de energía, para las “ondas” de probabilidad que se mueven a través del cristal con la velocidad de grupo de la ecuación (13.18), para la masa efectiva, etc. Solamente que ahora las ondas describen el comportamiento del *electrón que falta*—o “hueco” como es llamado. Así, un “hueco” se comporta del mismo modo que una partícula con una cierta masa efectiva m_{ef} . Pueden ver que esta partícula parecerá tener una carga positiva. Haremos énfasis sobre esos huecos en el próximo capítulo.

Como otro ejemplo, podemos considerar una línea de átomos *neutros* idénticos, uno de los cuales está en un estado excitado—esto es, con más energía que la de su estado fundamental. Digamos que C_n es la amplitud para que el átomo n -ésimo sea el excitado. Puede interactuar con un átomo vecino, pasándole a éste la energía en exceso y volviendo al estado fundamental. Llamemos a la amplitud para este proceso $iA/\hbar$. Se darán cuenta que se repiten las mismas matemáticas. Ahora el objeto que se mueve se le llama un *excitón*. Se comporta como una “partícula” neutra moviéndose a través del cristal, transportando la energía de excitación. Este movimiento puede estar implicado en ciertos procesos biológicos como la visión, o la fotosíntesis. Se ha supuesto que la absorción de la luz en la retina produce un “excitón” que se mueve a través de alguna estructura periódica (como las capas en las barras que describimos en el capítulo 36 del volumen I, figura 36-5) para acumularse en alguna estación especial, donde se usa la energía para inducir una reacción química.

13-6 Dispersión por imperfecciones en la red

Ahora queremos considerar el caso de un solo electrón en un cristal que no es perfecto. Nuestro análisis anterior dice que los cristales perfectos tienen conductividad perfecta—que los electrones pueden deslizarse por el cristal, como en el vacío, sin fricción. Una de las cosas más importantes que puede impedir que el electrón continúe así para siempre es una irregularidad o imperfección del cristal. Supongan, por ejemplo, que en alguna parte del cristal falta un átomo; o que alguien puso un átomo equivocado en uno de los lugares atómicos, de tal manera que las cosas ahí son diferentes a los otros lugares atómicos. Digamos que la energía E_0 o la amplitud A podían ser diferentes. ¿Cómo describiríamos entonces lo que sucede?

Para ser concretos, volveremos al caso unidimensional y supondremos que el átomo número “cero” es un átomo de “impureza”, y que tiene un valor diferente de E_0 que todos los otros átomos. Llamemos a esta energía $(E_0 + F)$. ¿Qué sucede? Cuando un electrón llega al átomo “cero”, hay cierta probabilidad para

que el electrón sea dispersado hacia atrás. Si un paquete de ondas se mueve a lo largo de la línea y llega a un lugar donde las cosas son algo diferentes, una parte de él continuará hacia adelante y otra parte rebotará hacia atrás. Es bastante difícil analizar esta situación usando un paquete de ondas, porque todo varía con el tiempo. Es mucho más fácil trabajar con soluciones estacionarias. Por ello trabajaremos con estados estacionarios, los cuales encontraremos que se pueden construir con ondas continuas que poseen partes transmitidas y reflejadas. En tres dimensiones, llamaríamos a la parte reflejada onda dispersada, puesto que se esparciría en diversas direcciones.

Empezamos con un conjunto de ecuaciones que son exactamente como las ecuaciones (13.6), a excepción de que la ecuación para $n = 0$ es diferente del resto. Las cinco ecuaciones para $n = -2, -1, 0, +1$ y $+2$ son como:

$$\begin{aligned} & \vdots & & \vdots \\ Ea_{-2} &= E_0 a_{-2} - Aa_{-1} - Aa_{-3}, \\ Ea_{-1} &= E_0 a_{-1} - Aa_0 - Aa_{-2}, \\ Ea_0 &= (E_0 + F)a_0 - Aa_1 - Aa_{-1}, \\ Ea_1 &= E_0 a_1 - Aa_2 - Aa_0, \\ Ea_2 &= E_0 a_2 - Aa_3 - Aa_1, \\ & \vdots & & \vdots \end{aligned} \tag{13.28}$$

Naturalmente, están todas las otras ecuaciones para las cuales $|n|$ es mayor que 2. Se parecerán a la ecuación (13.6).

Para el caso general, en realidad deberíamos usar una A diferente para la amplitud para que el electrón salte desde o hacia el átomo “cero”, pero las principales características de lo que ocurre se obtendrán de un ejemplo simplificado, en el que todas las A son iguales.

La ecuación (13.10) aún servirá como solución para todas las ecuaciones, excepto para la del átomo “cero”—no es apropiada para esta ecuación. Necesitamos una solución diferente que podemos construir del modo siguiente. La ecuación (13.10) representa a una onda que avanza en la dirección x positiva. Una onda que avance en la dirección x negativa habría sido una solución igualmente buena. Se hubiera escrito

$$a(x_n) = e^{-ikx_n}.$$

La solución más general que podíamos haber tomado para la ecuación (13.6) sería una combinación de una onda hacia adelante y otra hacia atrás, esto es

$$a_n = \alpha e^{ikx_n} + \beta e^{-ikx_n}. \tag{13.29}$$

Esta solución representa a una onda compleja de amplitud α , que se mueve en la dirección $+x$, y una onda de amplitud β , que se mueve en la dirección $-x$.

Ahora, observen el conjunto de ecuaciones para nuestro nuevo problema—las de (13.28) junto con las ecuaciones para todos los otros átomos. Las ecuaciones que incluyen las a_n con $n \leq -1$ se satisfacen con la ecuación (13.29), con la condición que k esté relacionada con E y con el espaciamiento de la red b mediante

$$E = E_0 - 2A \cos kb. \tag{13.30}$$

El resultado físico es una onda “incidente” de amplitud α que se acerca al átomo “cero” (el “dispersor”) por la izquierda, y una onda “dispersada” o “reflejada” de amplitud β que vuelve hacia la izquierda. No perdemos generalidad si establecemos que la amplitud α de la onda incidente igual a 1. La amplitud β es, en general, un número complejo.

Podemos decir exactamente lo mismo para las soluciones de a_n cuando $n \geq 1$. Los coeficientes podrían ser diferentes, por lo que tendríamos

$$a_n = \gamma e^{ikx_n} + \delta e^{-ikx_n}, \quad \text{para} \quad n \geq 1. \tag{13.31}$$

Aquí, γ es la amplitud de la onda que va hacia la derecha y δ una onda que viene desde la derecha. Queremos considerar la situación *física* en la que una onda viene originalmente sólo de la izquierda, y hay solamente una onda “transmitida” que sale al otro lado del dispersor—o átomo de “impureza”. Probaremos una solución en la que $\delta = 0$. Ciertamente podemos satisfacer todas las ecuaciones para las a_n , excepto por las tres de en medio en la ecuación (13.28), mediante las siguientes soluciones de prueba.

$$\begin{aligned} a_n \text{ (para } n < 0) &= e^{ikx_n} + \beta e^{-ikx_n}, \\ a_n \text{ (para } n > 0) &= \gamma e^{ikx_n}. \end{aligned} \quad (13.32)$$

la situación de la que estamos hablando se ilustra en la figura 13-6.

Usando las fórmulas en la ecuación (13.32) para a_{-1} y a_{+1} , las tres ecuaciones de en medio en (13.28) nos permitirán despejar a_0 , y también los dos coeficientes β y γ . Por lo tanto, hemos encontrado una solución completa. Diciendo que $x_n = nb$, tenemos que resolver las tres ecuaciones

$$\begin{aligned} (E - E_0)\{e^{ik(-b)} + \beta e^{-ik(-b)}\} &= -A\{a_0 + e^{ik(-2b)} + \beta e^{-ik(-2b)}\}, \\ (E - E_0 - F)a_0 &= -A\{\gamma e^{ikb} + e^{ik(-b)} + \beta e^{-ik(-b)}\}, \\ (E - E_0)\gamma e^{ikb} &= -A\{\gamma e^{ik(2b)} + a_0\}. \end{aligned} \quad (13.33)$$

Recuerden que E está dado en función de k por la ecuación (13.30). Si sustituyen este valor de E en las ecuaciones, y recuerdan que $\cos x = \frac{1}{2}(e^{ix} + e^{-ix})$, obtienen de la primera ecuación

$$a_0 = 1 + \beta, \quad (13.34)$$

y de la tercera ecuación

$$a_0 = \gamma. \quad (13.35)$$

Éstas dos son consistentes solamente si

$$\gamma = 1 + \beta. \quad (13.36)$$

Esta ecuación nos dice que la onda transmitida (γ) es justamente la onda incidente original (1) con una onda adicional (β), igual a la onda reflejada. Esto no siempre es verdad, pero resulta que sí lo es para una dispersión en un solo átomo. Si hubiese un grupo de átomos de impureza, la cantidad añadida a la onda que avanza no sería necesariamente igual a la onda reflejada.

Podemos obtener la amplitud β de la onda reflejada a partir de la ecuación de en medio de (13.33); encontramos que

$$\beta = \frac{-F}{F - 2iA \sin kb}. \quad (13.37)$$

Tenemos la solución completa para una red con un átomo extraño.

Se estarán preguntando por qué la onda transmitida puede ser “más” que la onda incidente, como nos muestra la ecuación (13.34). Recuerden, sin embargo, que β y γ son números complejos y que el número de partículas (o mejor, la probabilidad para encontrar una partícula) en una onda es proporcional al cuadrado de la norma de la amplitud. De hecho, habrá “conservación de electrones” solamente si

$$|\beta|^2 + |\gamma|^2 = 1. \quad (13.38)$$

Pueden demostrar que esto es verdad para nuestra solución.

13-7 Atrampar con una imperfección en la red

Hay otra cuestión interesante que puede surgir si F es un número negativo. Si la energía del electrón es más baja en el átomo de impureza (para $n = 0$) que

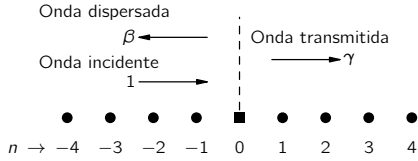


Figura 13-6. Ondas en una red unidimensional que tiene un átomo de “impureza” en $n = 0$.

en cualquier otro lugar, entonces el electrón puede quedar atrapado en este átomo. Esto es, si $(E_0 + F)$ está por debajo del mínimo de la banda $(E_0 - 2A)$, entonces el electrón puede ser “atrapado” en un estado con $E < E_0 - 2A$. Tal solución se puede obtener de lo que hasta ahora hemos hecho. Podemos obtener la solución, sin embargo, si permitimos que la solución de prueba que usamos en la ecuación (13.10) tenga un k imaginario. Hagamos que $k = \pm i\kappa$. De nuevo, podemos tener diferentes soluciones para $n < 0$ y para $n > 0$. Una solución posible para $n < 0$ podría ser

$$a_n \text{ (para } n < 0) = ce^{+\kappa x_n}. \quad (13.39)$$

Tenemos que tomar el signo positivo en el exponente; de lo contrario, la amplitud aumentaría indefinidamente para valores negativos grandes de n . Del mismo modo, una solución posible para $n > 0$ sería

$$a_n \text{ (para } n > 0) = c'e^{-\kappa x_n}. \quad (13.40)$$

Si sustituimos estas soluciones de prueba en (13.28) se satisfacen todas las ecuaciones, excepto las tres del medio, siempre que

$$E = E_0 - A(e^{\kappa b} + e^{-\kappa b}). \quad (13.41)$$

Ya que la suma de los dos términos exponenciales es siempre mayor que 2, esta energía está por debajo de la banda regular, que es lo que estamos buscando. Las tres ecuaciones restantes en (13.28) se satisfacen si $a_0 = c = c'$ y si κ se escoge en forma tal que

$$A(e^{\kappa b} - e^{-\kappa b}) = -F. \quad (13.42)$$

Combinando esta ecuación con la (13.41), podemos encontrar la energía del electrón atrapado; obtenemos

$$E = E_0 - \sqrt{4A^2 + F^2}. \quad (13.43)$$

El electrón atrapado tiene una energía única—ubicada por debajo de la banda de conducción.

Noten que las amplitudes, que tenemos en las ecuaciones (13.39) y (13.40), *no nos dicen* que el electrón atrapado está ubicado en el átomo de impureza. La probabilidad para encontrar al electrón en átomos vecinos está dada por el cuadrado de estas amplitudes. Para una elección particular de los parámetros, podría variar como se muestra en la gráfica de barras de la figura 13-7. La probabilidad es mayor para encontrar el electrón en el átomo de impureza. Para átomos cercanos, la probabilidad decae exponencialmente en función de la distancia al átomo de impureza. Éste es otro ejemplo de “penetración de barrera”. Desde el punto de vista de la física clásica, el electrón no tiene energía suficiente para alejarse del “pozo” de energía en el centro captor. Pero de manera cuántica, se puede alejar una pequeña distancia.

13-8 Amplitudes de dispersión y estados enlazados

Finalmente, podemos utilizar nuestro ejemplo para ilustrar un punto que es muy útil hoy en día en la física de las partículas de alta energía. Tiene que ver con una relación entre las amplitudes de dispersión y los estados enlazados. Supongan que hemos descubierto —a través de experimentos y análisis teórico— el modo en que los piones se dispersados por los protones. Entonces se descubre una nueva partícula, y alguien se pregunta si quizá sea solamente una combinación de un pión y un protón unidos en algún estado enlazado (en analogía a la manera en que un electrón está enlazado a un protón para formar un átomo de hidrógeno). Por estado enlazado, entendemos una combinación que tiene una energía más baja que la correspondiente a las dos partículas libres.

Hay una teoría general que dice que un estado enlazado existirá en aquella energía para la cual la amplitud de dispersión se hace infinita si se extrapola

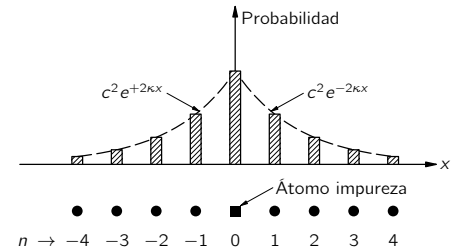


Figura 13-7. Las probabilidades relativas para encontrar un electrón atrapado en las posiciones atómicas cercanas al átomo de impureza.

algebraicamente (el término matemático es “se continúa analíticamente”) a las regiones de energía fuera de la banda permitida.

La razón física de esto es la siguiente. Un estado enlazado es una situación en la cual hay solamente ondas atadas a un punto, y no hay ninguna onda que llegue para generarlas, y simplemente existen ahí por sí mismas. La proporción relativa entre la llamada onda “dispersada” o creada y la onda que ha sido “enviada” es infinita. Podemos comprobar esta idea en nuestro ejemplo. Escribamos nuestra expresión (13.37) para la amplitud de dispersión directamente en términos de la energía E de la partícula que ha sido dispersada (en lugar de en términos de k). Ya que la ecuación (13.30) se puede escribir en la forma

$$2A \sin kb = \sqrt{4A^2 - (E - E_0)^2},$$

la amplitud de dispersión es

$$\beta = \frac{-F}{F - i\sqrt{4A^2 - (E - E_0)^2}}. \quad (13.44)$$

De acuerdo con nuestro procedimiento, esta ecuación se debería usar solamente para estados reales—los cuales tienen energías dentro de la banda de energía, $E = E_0 \pm 2A$. Pero supongan que olvidamos este hecho y extendemos la fórmula a regiones de energías “no físicas”, donde $|E - E_0| > 2A$. Para estas regiones no físicas podemos escribir[‡]

$$\sqrt{4A^2 - (E - E_0)^2} = i\sqrt{(E - E_0)^2 - 4A^2}.$$

Entonces la “amplitud de dispersión”, signifique lo que signifique, es

$$\beta = \frac{-F}{F + \sqrt{(E - E_0)^2 - 4A^2}}. \quad (13.45)$$

Ahora nos preguntamos: ¿Existe alguna energía E para la cual β se haga infinita (por ejemplo, para la cual la expresión para β tenga un “polo”)? Sí, siempre que F sea negativa, el denominador de (13.45) será cero cuando

$$(E - E_0)^2 - 4A^2 = F^2,$$

o cuando

$$E = E_0 \pm \sqrt{4A^2 + F^2}.$$

El signo menos da justamente la energía que encontramos en la ecuación (13.43) para la energía de atrampe.

¿Qué ocurre con el signo positivo? Éste nos da una energía *por encima* de la banda de energía permitida. Y verdaderamente ahí hay otro estado enlazado, que se nos escapó cuando resolvimos las ecuaciones (13.28). Les dejamos como un acertijo encontrar las energías y las amplitudes a_n para este estado enlazado.

La relación entre estados de dispersión y estados ligados suministra una de las pistas más útiles en la investigación actual sobre la comprensión de las observaciones experimentales acerca de las nuevas partículas extrañas.

[‡] El signo de la raíz a escoger aquí es un punto técnico, relacionado a los signos permitidos para κ en las ecuaciones (13.39) y (13.40). No lo trataremos aquí.

Semiconductores

14-1 Electrones y huecos en semiconductores

Uno de los notables y sorprendentes desarrollos recientes ha sido la aplicación de la ciencia del estado sólido al desarrollo técnico de dispositivos eléctricos, tales como los transistores. El estudio de los semiconductores llevó al descubrimiento de sus propiedades útiles y a una gran cantidad de aplicaciones prácticas. Este campo está cambiando tan rápidamente que lo que digamos hoy puede que sea incorrecto el año próximo. Por lo menos será incompleto. Y es perfectamente claro que, con el continuo estudio de estos materiales, serán posibles muchas cosas nuevas y maravillosas a medida que pase el tiempo. No necesitan entender este capítulo para comprender el resto de este volumen, pero es interesante darse cuenta que una parte de lo que están aprendiendo tiene cierta relación con el mundo práctico.

Hay una gran cantidad de semiconductores conocidos, pero nos concentraremos en los que actualmente tienen mayor aplicación técnica. También son los que se comprenden mejor, y al entenderlos, alcanzaremos cierto conocimiento de muchos otros. Las sustancias semiconductoras que hoy se usan más comúnmente son el silicio y el germanio. Estos elementos cristalizan en una red como la del diamante, un tipo de estructura cúbica, en la que los átomos tienen enlaces tetraédricos con sus cuatro vecinos más cercanos. Son aislantes a temperaturas muy bajas —cerca del cero absoluto— pero sí conducen algo de electricidad a temperatura ambiente. No son metales; se denominan *semiconductores*.

Si, de algún modo, ponemos un electrón adicional en un cristal de silicio o de germanio a baja temperatura, tendremos justamente la situación que describimos en el capítulo anterior. El electrón podrá vagar por el cristal, saltando de una posición atómica a la siguiente. En realidad sólo hemos examinado el comportamiento de los electrones en una red rectangular, y las ecuaciones serían un poco diferentes para la red de silicio o de germanio real. Sin embargo, todos los puntos esenciales están ilustrados por los resultados correspondientes de la red rectangular.

Como vimos en el capítulo 13, estos electrones sólo pueden tener energías dentro de cierta banda de energía —llamada *banda de conducción*—. En esta banda, la energía está relacionada con el número de onda $\mathbf{k}$ de la amplitud de probabilidad C [véase la ecuación (13.24)] mediante

$$E = E_0 - 2A_x \cos k_x a - 2A_y \cos k_y b - 2A_z \cos k_z c. \quad (14.1)$$

Las A son las amplitudes para saltar en las direcciones x , y y z , mientras que a , b y c son los espaciamentos de la red en dichas direcciones.

Para energías cercanas al mínimo de la banda, podemos aproximar la ecuación (14.1) con

$$E = E_{\min} + A_x a^2 k_x^2 + A_y b^2 k_y^2 + A_z c^2 k_z^2 \quad (14.2)$$

(véase la sección 13-4).

Si consideramos el movimiento del electrón en una dirección determinada, de modo que las componentes de $\mathbf{k}$ siempre estén en la misma proporción, la energía es una función cuadrática del número de onda —y, como hemos visto, del momentum del electrón—. Podemos escribir

$$E = E_{\min} + \alpha k^2, \quad (14.3)$$

14-1 Electrones y huecos en semiconductores

14-2 Semiconductores impuros

14-3 El efecto Hall

14-4 Junturas de semiconductores

14-5 Rectificación en una juntura de semiconductores

14-6 El transistor

Referencia: C. Kittel, *Introduction to Solid State Physics*, 2.^a ed., John Wiley and Sons, Inc., Nueva York, 1956. Capítulos 13, 14 y 18.

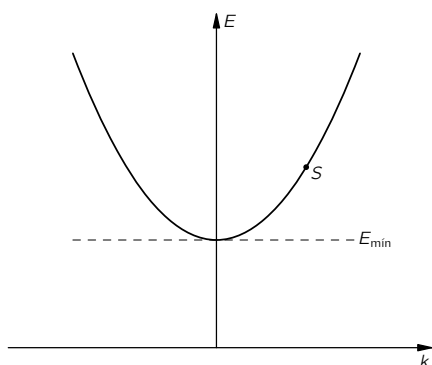


Figura 14-1. El diagrama de energía para un electrón en un cristal aislante.

donde α es una constante; podemos hacer una gráfica de E en función de k como el de la figura 14-1. Llamaremos a esta gráfica “diagrama de energía”. Un electrón en un estado determinado de energía y de momentum se puede indicar con un punto, como S , en la figura.

Como mencionamos en el capítulo 13, podemos tener una situación similar si quitamos un electrón de un aislante neutro. Así, un electrón de un átomo cercano puede saltar y llenar el “hueco”, pero dejando otro “hueco” en el átomo de donde partió. Podemos describir este comportamiento escribiendo una amplitud para encontrar al hueco en cualquier átomo determinado, y diciendo que el hueco puede saltar de un átomo al otro. (Es claro que la amplitud A para que el hueco salte del átomo a al átomo b es exactamente la misma que la amplitud para que un electrón que está en el átomo b salte dentro del hueco en el átomo a .) Las matemáticas son las mismas para tratar al hueco que las que teníamos para el electrón adicional; obtenemos de nuevo que la energía del hueco está relacionada con su número de onda mediante una ecuación como la (14.1) o la (14.2), excepto, naturalmente, con valores numéricos diferentes para las amplitudes A_x , A_y y A_z . El hueco tiene una energía relacionada con el número de onda de sus amplitudes de probabilidad. Su energía está dentro de una banda restringida; cerca del fondo de la banda, su energía varía cuadráticamente con el número de onda —o el momentum— como en la figura 14-1. Siguiendo los argumentos de la sección 13-3, encontramos que el hueco también se comporta como una partícula clásica con una cierta masa efectiva —excepto que en cristales no cúbicos, la masa depende de la dirección de movimiento—. Así entonces, el hueco se comporta como una partícula positiva moviéndose a través del cristal. La carga del hueco-partícula es positiva, porque está ubicada en el sitio donde falta un electrón; cuando se mueve en una dirección hay, en realidad, electrones moviéndose en la dirección opuesta.

Si ponemos varios electrones en un cristal neutro, éstos se moverán como los átomos de un gas a baja presión. Si no hay muchos, sus interacciones no serán muy importantes. Si después aplicamos un campo eléctrico al cristal, los electrones comenzarán a moverse y circulará una corriente eléctrica. Eventualmente, éstos serán arrastrados hasta uno de los bordes del cristal, y, si ahí hay un electrodo de metal, serán recogidos, dejando neutro el cristal.

Análogamente, podríamos poner muchos huecos en un cristal. Se moverían a través de él al azar, a no ser que hubiera un campo eléctrico. En un campo, éstos circularían hacia la terminal negativa, y serían “recogidos” —lo que en realidad ocurre es que se neutralizan con los electrones de la terminal metálica—.

Además, se puede tener electrones y huecos al mismo tiempo. Si no son demasiados, todos seguirán su camino independientemente. En un campo eléctrico, todos contribuirán a la corriente. Por razones obvias, los electrones se denominan *portadores negativos* y los huecos *portadores positivos*.

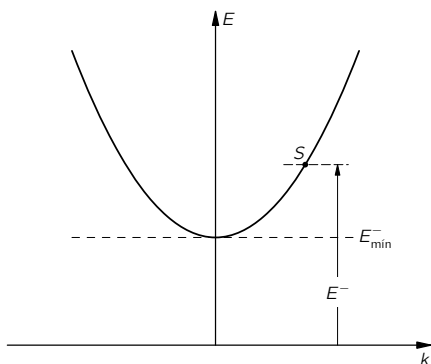


Figura 14-2. Se necesita la energía E^- para “crear” un electrón libre.

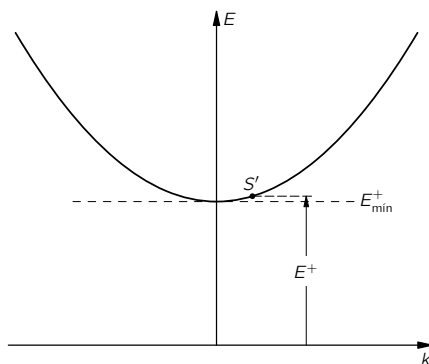


Figura 14-3. Se necesita la energía E^+ para “crear” un hueco en el estado S' .

Hasta ahora, hemos considerado que los electrones se introducen en el cristal desde afuera, o se quitan para hacer un hueco. También es posible “crear” un par electrón-hueco, sacando un electrón enlazado de un átomo neutro y poniéndolo a cierta distancia en el mismo cristal. Así tenemos un electrón libre y un hueco libre, y los dos se pueden mover como hemos descrito.

La energía necesaria para poner un electrón en un estado S —decimos para “crear” el estado S — es la energía E^- que muestra la figura 14-2. Es una energía por encima de $E_{\min}^-$. La energía necesaria para “crear” un hueco en un estado S' es la energía E^+ de la figura 14-3, la cual es una energía mayor que $E_{\min}^+$. Ahora bien, si creamos un par en los estados S y S' , la energía necesaria es simplemente $E^- + E^+$.

La creación de pares es un proceso común (como veremos más adelante), y por ello, a mucha gente le gusta mostrar a las figuras 14-2 y 14-3 en la misma gráfica —con la energía del hueco graficada *hacia abajo*, aunque, naturalmente, es una energía *positiva*—. Hemos combinado nuestras dos gráficas en la figura 14-4. La ventaja de esta gráfica es que la energía $E_{\text{par}} = E^- + E^+$ necesaria para crear un par, con el electrón en S y el hueco en S' , es simplemente la distancia vertical entre S y S' , como se muestra en la figura 14-4. La mínima energía necesaria para crear un par se denomina “diferencia” de energía, y es igual a $E_{\min}^- + E_{\min}^+$.

Algunas veces verán un diagrama más simple, llamado diagrama de niveles de energía, que trazamos cuando no estamos interesados en la variable k . Este diagrama —mostrado en la figura 14-5— sólo muestra las energías posibles para los electrones y los huecos.†

¿Cómo se puede crear pares de electrón-hueco? Hay varias maneras. Por ejemplo, se pueden absorber fotones de luz (o de rayos X) y crear un par, si la energía del fotón es mayor que la diferencia de energía. La rapidez con la que se producen los pares es proporcional a la intensidad de la luz. Si se montan dos electrodos sobre un disco del cristal, y se aplica un voltaje de “polarización”, los electrones y los huecos serán arrastrados hacia los electrodos. La corriente del circuito será proporcional a la intensidad de la luz. Este mecanismo es responsable por el fenómeno de la fotoconductividad y del funcionamiento de las celdas fotoconductoras.

Los pares electrón-hueco también se pueden producir con partículas de alta energía. Cuando una partícula cargada se mueve rápido —por ejemplo, un protón, o un pion, con una energía de decenas o centenares de MeV— y atraviesa un cristal, su campo eléctrico arrancará electrones de sus estados enlazados, creando pares de electrón-hueco. Estos eventos ocurren cientos de miles de veces por milímetro de camino. Después del paso de la partícula, se puede recoger a los portadores, obteniéndose un pulso eléctrico. Éste es el mecanismo en juego en los contadores hechos con semiconductores, los cuales se han puesto recientemente en uso en la física nuclear experimental. Esos contadores no necesitan semiconductores, también se pueden hacer con aislantes cristalinos. En realidad, el primero de esos contadores se hizo con un cristal de diamante, que es aislante a temperatura ambiente. Son necesarios cristales muy puros para que electrones y huecos puedan moverse libremente hacia los electrodos sin ser atrapados. Se utilizan los semiconductores de silicio y de germanio, porque éstos se pueden producir con gran pureza y en tamaños razonablemente grandes (del orden de centímetros).

Hasta ahora, nos hemos ocupado de cristales semiconductores a temperaturas cercanas al cero absoluto. A cualquier temperatura finita, hay aún otro mecanismo para crear pares de electrón-hueco. La energía térmica del cristal puede suministrar la energía del par. Las vibraciones térmicas del cristal pueden transferir su energía a un par —dando lugar a la creación “espontánea”—.

† En muchos libros, este mismo diagrama se interpreta de manera diferente. La escala de energía sólo se refiere a los *electrones*. En lugar de considerar la energía del hueco, consideran la energía que *tendría* un electrón si llenara el hueco. Esta energía es *menor* que la energía de un electrón libre —de hecho, es más baja justamente en la cantidad que ven en la figura 14-5—. Con esta interpretación de la escala de energía, la diferencia de energía es la mínima energía que se le debe dar a *un electrón* para que pase de su estado enlazado a la banda de conducción.

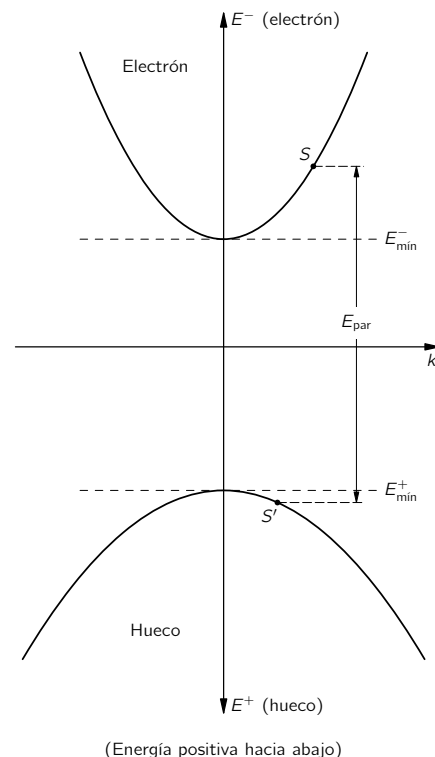


Figura 14-4. Diagramas de energía para un electrón y un hueco, graficadas juntas.

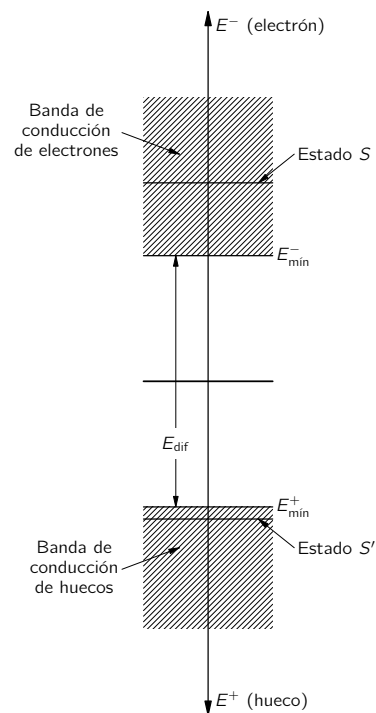


Figura 14-5. Diagrama de niveles de energía para electrones y huecos.

La probabilidad por unidad de tiempo para que una energía igual a la diferencia de energía E_{dif} se concentre en la posición de un sitio atómico, es proporcional a $e^{-E_{\text{dif}}/\kappa T}$, donde T es la temperatura y κ es la constante de Boltzmann (véase el capítulo 40 del volumen I). Cerca del cero absoluto no hay una probabilidad apreciable, pero conforme aumenta la temperatura, hay una probabilidad creciente para producir esos pares. A cualquier temperatura finita, la producción debería continuar para siempre, a una rapidez constante, produciendo más y más portadores negativos y positivos. Por supuesto, eso no sucede porque, después de cierto tiempo, los electrones y los huecos se encuentran accidentalmente —el electrón cae en el hueco y la energía en exceso se transfiere a la red—. Decimos que el electrón y el hueco “se aniquilan”. Hay cierta probabilidad por segundo para que un hueco encuentre a un electrón y que los dos se aniquilen mutuamente.

Si el número de electrones por unidad de volumen es N_n (n indica portadores negativos) y la densidad de portadores positivos es N_p , la probabilidad por unidad de tiempo para que un electrón y un hueco se encuentren y se aniquilen es proporcional al producto $N_n N_p$. En el equilibrio, esta rapidez debe ser igual a la rapidez con que se crean pares. Pueden ver que, en el equilibrio, el producto de N_n y N_p debe estar dado por una constante multiplicada por el factor de Boltzmann:

$$N_n N_p = (\text{const}) e^{-E_{\text{dif}}/\kappa T}. \quad (14.4)$$

Al decir constante, queremos decir casi constante. Una teoría más completa —que incluye más detalles de cómo los huecos y los electrones “se encuentran”— muestra que la “constante” depende ligeramente de la temperatura, pero que la mayor dependencia de la temperatura está en el exponencial.

Como ejemplo, consideremos un material puro y originalmente neutro. A una temperatura finita, esperarían que el número de portadores negativos fuera igual al de positivos, $N_n = N_p$. Por lo tanto, cada uno de ellos debería variar con la temperatura en la forma $e^{-E_{\text{dif}}/2\kappa T}$. La variación de muchas de las propiedades de un semiconductor —la conductividad, por ejemplo— está determinada principalmente por el factor exponencial, porque todos los otros factores varían mucho más lentamente con la temperatura. La diferencia de energía del germanio es de unos 0.72 eV y la del silicio 1.1 eV.

A temperatura ambiente, κT es aproximadamente 1/40 de un electronvoltio. A estas temperaturas hay suficientes huecos y electrones como para dar una conductividad apreciable, mientras que a, digamos, 30K —un décimo de la temperatura ambiente— la conductividad es imperceptible. La diferencia de energía del diamante es de 6 o 7 eV, y el diamante es un buen aislante a la temperatura ambiente.

14-2 Semiconductores impuros

Hasta ahora, hemos hablado de dos maneras para poner electrones adicionales en una red cristalina, la cual, por lo demás, es idealmente perfecta. Una manera era inyectar el electrón desde una fuente externa; la otra era arrancar un electrón enlazado de un átomo neutro, creando simultáneamente un electrón y un hueco. Es posible poner electrones en la banda de conducción de un cristal aun de otra manera. Supongan que tenemos un cristal de germanio en el cual uno de los átomos de germanio es remplazado por un átomo de arsénico. Los átomos de germanio tienen valencia 4, y la estructura del cristal es controlada por los cuatro electrones de valencia. Por otra parte, el arsénico tiene valencia 5. Ocurre que un solo átomo de arsénico puede instalarse en la red del germanio (porque tiene aproximadamente el tamaño apropiado), pero, para hacerlo, debe actuar como un átomo de valencia 4 —usa cuatro de sus electrones de valencia para formar los enlaces cristalinos y le sobra un electrón—. Este electrón adicional está unido muy débilmente —la energía de enlace es casi 1/100 de un electronvoltio—. A temperatura ambiente, el electrón puede tomar fácilmente esa energía de la energía térmica del cristal e irse por su propia cuenta —moviéndose por la red como un electrón libre—. Un átomo de impureza, como el arsénico, se denomina

sitio donador, porque puede dar un portador negativo al cristal. Si se genera un cristal de germanio a partir de la fase fundida, a la que se le ha agregado una cantidad muy pequeña de arsénico, los donadores de arsénico estarán distribuidos por todo el cristal; el cristal tendrá una cierta densidad de portadores negativos incluida en su estructura.

Ustedes podrían pensar que estos portadores serían barridos por completo, cuando se aplica algún campo eléctrico pequeño al cristal. Sin embargo, esto no sucede, porque los átomos de arsénico incluidos en el cuerpo del cristal tienen cargas positivas. Para que el cristal permanezca neutro, la densidad promedio de los electrones portadores negativos debe ser igual a la de donadores. Si se ponen dos electrodos en los bordes de ese cristal y se conectan a una batería, circulará corriente; pero, a medida que se extraen los electrones portadores por un extremo, se deben introducir nuevos electrones de conducción por el otro extremo, de manera que la densidad promedio de electrones de conducción sea prácticamente igual a la densidad de sitios donadores.

Como los sitios donadores están cargados positivamente, éstos tendrán cierta tendencia a capturar algunos electrones de conducción a medida que se difunden por el cristal. En consecuencia, un donador puede actuar como una trampa, como las que mencionamos en la sección anterior. Pero si la energía de atrape es suficientemente pequeña —como en el caso del arsénico— el número de portadores que están atrapados en cualquier instante es una fracción pequeña del total. Para comprender completamente el comportamiento de los semiconductores, se debe tener en cuenta este atrape. No obstante, en el resto de nuestra discusión, supondremos que la energía de atrape es suficientemente baja, y que la temperatura es suficientemente alta, como para que todos los sitios donadores hayan liberado a sus electrones. Por supuesto, esto es sólo una aproximación.

También sería posible incluir en un cristal de germanio un átomo de impureza cuya valencia sea 3, como el aluminio. El átomo de aluminio trata de actuar como si tuviera valencia 4 robando un electrón adicional. Puede robar un electrón de algún átomo vecino de germanio, y convertirse en un átomo cargado negativamente, con una valencia efectiva de 4. Sin duda, cuando roba un electrón de un átomo de germanio, deja un hueco en él; este hueco puede vagar por el cristal como portador positivo. Un átomo de impureza que puede producir un hueco así se denomina *receptor*, porque “recibe” un electrón. Si se genera un cristal de germanio o de silicio a partir de una fase fundida, a la cual se le ha agregado una pequeña cantidad de impurezas de aluminio, el cristal tendrá incluida una cierta densidad de huecos que pueden actuar como portadores positivos.

Cuando se agrega una impureza donadora o receptora a un semiconductor, decimos que el cristal está “dopado”.

Cuando un cristal de germanio, con algunas impurezas donadoras incluidas, está a temperatura ambiente, hay unos electrones de conducción provenientes de la creación de pares electrón-hueco inducida térmicamente, y otros que provienen de los donadores. Naturalmente, los electrones de ambas fuentes son equivalentes, y su número total N_n es el que interviene en los procesos estadísticos que llevan al equilibrio. Si la temperatura no es muy baja, el número de portadores negativos con que contribuyen los átomos de impureza donadores, es aproximadamente igual al número de átomos de impureza presente. En el equilibrio, la ecuación (14.4) aún debe ser válida; a una temperatura dada el producto $N_n N_p$ está determinado. Esto significa que si agregamos impurezas donadoras, lo cual aumenta N_n , el número N_p de portadores positivos tendrá que disminuir de modo tal que $N_n N_p$ no cambie. Si la concentración de impurezas es suficientemente alta, el número N_n de portadores negativos está determinado por los donadores, y es prácticamente independiente de la temperatura —toda la variación del factor exponencial es proporcionada por N_p , aunque éste es mucho menor que N_n —. En un cristal mayormente puro, con una pequeña concentración de impurezas donadoras, la mayoría de los portadores serán negativos; este material se denomina semiconductor “tipo n ”.

Si se agrega una impureza tipo receptor a la red cristalina, algunos de los nuevos huecos se moverán, aniquilando algunos electrones libres producidos

por las fluctuaciones térmicas. Este proceso seguirá hasta que se satisfaga la ecuación (14.4). En condiciones de equilibrio, el número de portadores positivos aumentará y el de portadores negativos disminuirá, dejando constante al producto. Un material con un exceso de portadores positivos se denomina semiconductor “tipo p ”.

Si ponemos dos electrodos a un pedazo de cristal semiconductor y los conectamos a una fuente de diferencia de potencial, habrá un campo eléctrico dentro del cristal. El campo eléctrico hará que los portadores positivos y negativos se muevan, y circulará una corriente eléctrica. Consideremos primero lo que sucederá en un material tipo n , en el cual hay una gran mayoría de portadores negativos. En este material podemos ignorar a los huecos; éstos contribuirán muy poco a la corriente, porque hay muy pocos de ellos. En un cristal ideal, los portadores se moverían sin ningún impedimento. Sin embargo, en un cristal real a temperatura finita —especialmente en un cristal con algunas impurezas— los electrones no se mueven de manera completamente libre. Continuamente sufren colisiones que los sacan de sus trayectorias originales, o sea que están cambiando su momentum. Estas colisiones son justamente las dispersiones que discutimos en el capítulo anterior, y ocurren en cualquier irregularidad de la red cristalina. En un material tipo n , las causas principales de dispersión son precisamente los donadores que están produciendo los portadores. Como los electrones de conducción tienen una energía ligeramente diferente en los sitios donadores, las ondas de probabilidad se dispersan en esos puntos. Aun en un cristal perfectamente puro, hay (a cualquier temperatura finita) irregularidades en la red, debidas a las vibraciones térmicas. Desde el punto de vista clásico, podemos decir que los átomos no están alineados exactamente en una red regular, sino que, en cualquier instante, están ligeramente fuera de su sitio debido a sus vibraciones térmicas. La energía E_0 asociada a cada punto de la red en la teoría que hemos descrito en el capítulo 13, varía un poco de un punto a otro, por lo cual, las ondas de amplitud de probabilidad no se transmiten perfectamente, sino que se dispersan de manera irregular. A temperaturas muy altas, o para los materiales muy puros, esta dispersión puede llegar a ser importante, pero en la mayoría de los materiales dopados utilizados en las aplicaciones prácticas, los átomos de impureza producen la mayor parte de la dispersión. Ahora nos gustaría estimar la conductividad eléctrica de esta clase de materiales.

Cuando se aplica un campo eléctrico a un semiconductor tipo n , cada portador negativo se acelerará en este campo, aumentando su velocidad, hasta que se dispersa en uno de los donadores. Esto significa que los portadores que, por lo general, se mueven al azar con sus energías térmicas, adquirirán una velocidad promedio de arrastre en la dirección de las líneas del campo eléctrico, dando lugar a una corriente en el cristal. La velocidad de arrastre es, en general, bastante pequeña comparada con las velocidades térmicas típicas; por lo que podemos estimar la corriente, suponiendo que el tiempo promedio que el portador viaja entre dispersiones es constante. Digamos que el portador negativo tiene una carga eléctrica efectiva q_n . En un campo eléctrico $\mathcal{E}$, la fuerza que actúa sobre el portador será $q_n\mathcal{E}$. En la sección 43-3 del volumen I, calculamos la velocidad promedio de arrastre en estas circunstancias y encontramos que está dada por $F\tau/m$, donde F es la fuerza sobre la carga, τ es el tiempo libre medio entre colisiones y m es la masa. Deberíamos usar la masa efectiva que calculamos en el capítulo anterior, pero como sólo queremos hacer un cálculo aproximado, supondremos que esta masa efectiva es la misma en todas las direcciones. Aquí la llamaremos m_n . Con esta aproximación, la velocidad promedio de arrastre será

$$v_{\text{ar}} = \frac{q_n \mathcal{E} \tau_n}{m_n}. \quad (14.5)$$

Si conocemos la velocidad de arrastre, podemos encontrar la corriente. La densidad de corriente eléctrica, $\mathbf{j}$, es simplemente el número de portadores por unidad de volumen, N_n , multiplicado por la velocidad promedio de arrastre y por la carga de cada portador. Por lo tanto, la densidad de corriente es

$$\mathbf{j} = N_n \mathbf{v}_{ar} q_n = \frac{N_n q_n^2 \tau_n}{m_n} \mathbf{E}. \quad (14.6)$$

Vemos que la densidad de corriente es proporcional al campo eléctrico; ese material semiconductor sigue la ley de Ohm. El coeficiente de proporcionalidad entre $\mathbf{j}$ y $\mathbf{E}$, la conductividad σ , es

$$\sigma = \frac{N_n q_n^2 \tau_n}{m_n}. \quad (14.7)$$

En un material tipo n , la conductividad es relativamente independiente de la temperatura. En primer lugar, la cantidad de portadores N_n está determinada principalmente por la densidad de donadores en el cristal (mientras la temperatura no sea tan baja como para que muchos portadores sean atrapados). En segundo lugar, el tiempo medio τ_n entre colisiones está controlado principalmente por la densidad de átomos de impureza, que, por supuesto, es independiente de la temperatura.

Podemos aplicar los mismos argumentos para un material tipo p , cambiando únicamente los valores de los parámetros que aparecen en la ecuación (14.7). Si hay cantidades parecidas de portadores negativos y positivos presentes al mismo tiempo, debemos sumar las contribuciones de cada clase de portador. La conductividad total estará dada por

$$\sigma = \frac{N_n q_n^2 \tau_n}{m_n} + \frac{N_p q_p^2 \tau_p}{m_p}. \quad (14.8)$$

En los materiales muy puros N_p y N_n serán casi iguales. Éstos serán menores que en un material dopado, por lo que la conductividad será menor. Además, variarán rápidamente con la temperatura (en la forma $e^{-E_{dif}/2kT}$, como hemos visto), por lo cual la conductividad puede cambiar mucho muy rápido con la temperatura.

14-3 El efecto Hall

Es muy peculiar que en una sustancia, donde lo único relativamente libre son los electrones, haya una corriente eléctrica transportada por huecos que se comportan como partículas positivas. Por ello, deseamos describir un experimento, el cual muestra de manera bastante clara que el signo del portador de la corriente eléctrica es con toda seguridad positivo. Supongan que tenemos un bloque constituido con un material semiconductor —también podría ser un metal— y le aplicamos un campo eléctrico, de modo que haya una corriente en alguna dirección, por ejemplo, la dirección horizontal, como se muestra en la figura 14-6. Supongan ahora que aplicamos un campo magnético al bloque en dirección perpendicular a la corriente, digamos que *hacia* el plano de la figura. Los portadores en movimiento experimentarán una fuerza magnética $q(\mathbf{v} \times \mathbf{B})$. Y como la velocidad promedio de arrastre es hacia la derecha o hacia la izquierda —según sea el signo de la carga del portador— la fuerza magnética promedio sobre los portadores será hacia arriba o hacia abajo. No, ¡eso no es correcto! Para las direcciones que hemos supuesto para la corriente y el campo magnético, la fuerza magnética en las cargas en movimiento es siempre *hacia arriba*. Las cargas positivas que se mueven en la dirección de $\mathbf{j}$ (hacia la derecha) experimentan una fuerza hacia arriba. Si la corriente es transportada por cargas negativas, éstas se estarán moviendo hacia la izquierda (para el mismo signo de la corriente de conducción) y también experimentarán una fuerza hacia arriba. Sin embargo, en condiciones estacionarias, los portadores no se mueven hacia arriba, porque la corriente sólo puede circular de izquierda a derecha. Lo que ocurre es que al principio algunas cargas van hacia arriba, produciendo una densidad superficial de carga sobre la superficie superior del semiconductor —dejando una densidad superficial de carga igual y opuesta en la superficie inferior del cristal—. Las cargas se agrupan en las superficies de arriba y de abajo hasta que las fuerzas eléctricas, que producen sobre las cargas en movimiento, compensan exactamente

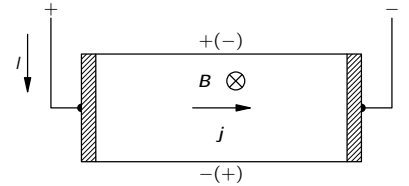


Figura 14-6. El efecto Hall proviene de las fuerzas magnéticas sobre los portadores.

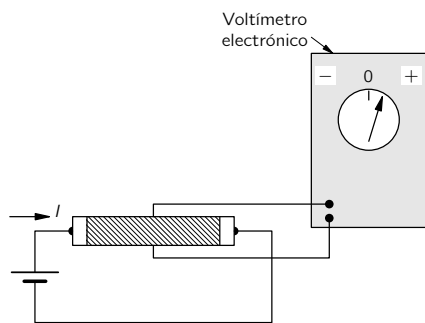


Fig. 14-7. Midiendo el efecto Hall.

la fuerza magnética (en promedio), de modo que la corriente estacionaria circula horizontalmente. Las cargas en las superficies de arriba y de abajo producirán una diferencia de potencial vertical en el cristal, la cual se puede medir con un voltímetro de alta resistencia, como lo muestra la figura 14-7. El signo de la diferencia de potencial registrada por el voltímetro depende del signo de las cargas portadoras responsables de la corriente.

Cuando estos experimentos se hicieron por primera vez, se esperaba que el signo de la diferencia de potencial fuera negativo, como sería de esperar para electrones de conducción negativos. De ahí que la gente se quedara muy sorprendida al encontrar que, para algunos materiales, el signo de la diferencia de potencial estuviera en la dirección opuesta. Parecía que el portador de la corriente era una partícula de carga positiva. De nuestra discusión de los semiconductores dopados, es comprensible que un semiconductor tipo n produzca el signo de la diferencia de potencial correspondiente a portadores negativos, y que un semiconductor tipo p dé una diferencia de potencial opuesta, ya que la corriente es transportada por huecos cargados positivamente.

El descubrimiento del signo anómalo de la diferencia de potencial en el efecto Hall fue originalmente en un metal; no en un semiconductor. Se había supuesto que en los metales, la conducción siempre se debía al electrón; sin embargo, se encontró que en el berilio, la diferencia de potencial tenía el signo que no correspondía. Hoy se entiende que en los metales, así como en los semiconductores, es posible que, en ciertas circunstancias, los “objetos” responsables de la conducción sean huecos. Aunque, en última instancia, sean los electrones los que se mueven, aun así, la relación entre el momentum y la energía y la respuesta a campos externos son exactamente las que se esperan para una corriente eléctrica transportada por partículas positivas.

Veamos si podemos hacer una estimación cuantitativa de la magnitud de la diferencia de potencial que se espera del efecto Hall. Si el voltímetro de la figura 14-7 extrae una corriente que podemos despreciar, entonces las cargas que están en el semiconductor se deben estar moviendo de izquierda a derecha, mientras que la fuerza magnética vertical debe estar exactamente compensada por un campo eléctrico vertical, que llamaremos $\mathcal{E}_{tr}$ (“tr” significa “transversal”). Si este campo eléctrico debe compensar las fuerzas magnéticas, debemos tener

$$\mathcal{E}_{tr} = -\mathbf{v}_{ar} \times \mathbf{B}. \quad (14.9)$$

Usando la relación entre la velocidad de arrastre y la densidad de corriente eléctrica dada en la ecuación (14.6), obtenemos

$$\mathcal{E}_{tr} = -\frac{1}{qN} jB.$$

La diferencia de potencial entre la parte superior y la base del cristal es, por supuesto, la intensidad de campo eléctrico multiplicada por la altura del cristal. La intensidad del campo eléctrico $\mathcal{E}_{tr}$ dentro del cristal es proporcional a la densidad de corriente y a la intensidad del campo magnético. La constante de proporcionalidad $1/qN$ se denomina coeficiente de Hall y comúnmente se lo representa con el símbolo R_H . El coeficiente de Hall depende solamente de la densidad de portadores —siempre que los portadores de un signo estén en gran mayoría—. Por tanto, la medición del efecto Hall es una manera conveniente para determinar experimentalmente la densidad de portadores en un semiconductor.

14-4 Junturas de semiconductores

Ahora queremos discutir lo que ocurre si tomamos dos pedazos de germanio o de silicio con características internas distintas —por ejemplo, con diferentes tipos o cantidades de dopaje— y los ponemos juntos para hacer una “juntura”. Comencemos con lo que se denomina juntura p - n , en la cual tenemos germanio tipo p de un lado de la frontera y germanio tipo n del otro lado de la frontera —como se muestra en la figura 14-8—. En realidad, no es práctico juntar dos

pedazos de cristal y tenerlos en contacto uniforme a la escala atómica. En su lugar, las juntas se hacen con un solo cristal modificado en las dos regiones. Una manera es agregar una impureza de dopaje apropiada a la fase “fundida” después que ha crecido solamente la mitad del cristal. Otra manera es pintar un poco la superficie con el elemento de impureza y luego calentar el cristal, haciendo que algunos átomos de impureza se difundan dentro del cristal. Las juntas hechas así no tienen una frontera definida, aunque las fronteras podrían tener un grosor de 10^{-4} centímetros aproximadamente. En nuestras discusiones imaginaremos una situación ideal, en la que estas dos regiones del cristal con propiedades diferentes se encuentran en una frontera bien definida.

En el lado tipo n de la junta p - n , hay electrones libres que se pueden mover por todas partes, y también sitios donadores fijos que balancean la carga total. En el lado tipo p , hay huecos libres moviéndose por todas partes, y un número igual de receptores negativos que compensan la carga. En realidad, ésta es la descripción de la situación antes de poner a los dos materiales en contacto. Una vez conectados, la situación cambia cerca de la frontera. Cuando los electrones del material tipo n llegan a la frontera, no son reflejados de vuelta, como ocurriría en una superficie libre, sino que pueden continuar hacia el material tipo p . Por lo tanto, algunos electrones del material tipo n tenderán a difundirse hacia el material tipo p , donde hay menos electrones. Esto no continúa para siempre porque, a medida que perdemos electrones del lado n , la carga positiva neta ahí aumenta, hasta que se forma un voltaje eléctrico, el cual retarda la difusión de los electrones hacia el lado p . De modo similar, los portadores positivos del material tipo p se pueden difundir a través de la junta hacia el material tipo n . Al hacerlo, dejan atrás un exceso de carga negativa. En condiciones de equilibrio, la corriente difusiva neta debe ser cero. Esto se debe a los campos eléctricos, los cuales se forman de tal modo, que arrastran a los portadores positivos de vuelta hacia el material tipo p .

Los dos procesos difusivos que hemos descrito ocurren simultáneamente y, como pueden observar, ambos actúan en la dirección en la que el material tipo n se carga positivamente y el tipo p se carga negativamente. Debido a la conductividad finita del material semiconductor, el cambio en el potencial entre el lado p y el lado n se producirá en una región estrecha, cerca de la frontera; el cuerpo principal de cada bloque de material tendrá potencial uniforme. Imaginemos un eje x en una dirección perpendicular a la superficie de la frontera. Entonces, el potencial eléctrico variará con x , como muestra la figura 14-9(b). También mostramos en la parte (c) de la figura, la variación esperada en la densidad N_n de portadores n y la de la densidad N_p de portadores p . Lejos de la junta, las densidades de portadores N_n y N_p deben ser simplemente la densidad de equilibrio, lo que se esperaría para bloques individuales de material a la misma temperatura. (Hemos trazado la figura para una junta en la que el material de tipo p está dopado más fuertemente que el material de tipo n .) Debido al gradiente de potencial en la junta, los portadores positivos tienen que subir un escalón de potencial para llegar al lado tipo n . Esto es, en condiciones de equilibrio, debe haber menos portadores positivos en el material tipo n que en el tipo p . Recordando las leyes de la mecánica estadística, esperamos que el cociente de portadores tipo p a ambos lados esté dado por la siguiente ecuación:

$$\frac{N_p(\text{lado } n)}{N_p(\text{lado } p)} = e^{-q_p V / \kappa T}. \quad (14.10)$$

El producto $q_p V$ en el numerador del exponente es justamente la energía necesaria para llevar una carga q_p a través de la diferencia de potencial V .

Tenemos una ecuación extremadamente parecida para las densidades de portadores tipo n :

$$\frac{N_n(\text{lado } n)}{N_n(\text{lado } p)} = e^{-q_n V / \kappa T}. \quad (14.11)$$

Si conocemos las densidades de equilibrio en los dos materiales, podemos utilizar cualquiera de las dos ecuaciones anteriores para determinar la diferencia de potencial en la junta.

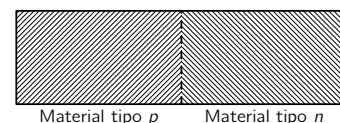


Figura 14-8. Una junta p - n .

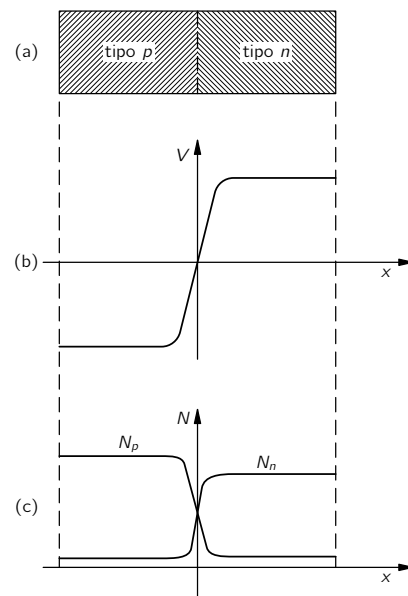


Figura 14-9. El potencial eléctrico y las densidades de portadores en una junta no polarizada de semiconductores.

Observen que si las ecuaciones (14.10) y (14.11) tienen que dar el mismo valor para la diferencia de potencial V , el producto $N_p N_n$ debe ser igual para el lado p y para el lado n . (Recuerden que $q_n = -q_p$.) No obstante, hemos visto antes que este producto sólo depende de la temperatura y de la diferencia de energía del cristal. Si ambos lados están a la misma temperatura, las dos ecuaciones son consistentes con el mismo valor de la diferencia de potencial.

Debido a que hay una diferencia de potencial de un lado de la juntura al otro, ésta se parece un poco a una batería. Tal vez si conectamos un alambre del lado tipo n al lado tipo p obtendremos una corriente eléctrica. Eso sería bueno, porque entonces la corriente circularía permanentemente, sin gastar ningún material, y tendríamos una fuente infinita de energía, ¡violando la segunda ley de la termodinámica! Pero no hay corriente si se conecta el lado p al lado n con un alambre. Y la razón es fácil de entender. Supongan que tenemos primero un alambre hecho con un trozo de material no dopado. Al conectar este alambre al lado tipo n tenemos una juntura. Habrá una diferencia de potencial en la juntura. Digamos que es justamente la mitad de la diferencia de potencial entre el material tipo p y el material tipo n . Al conectar nuestro alambre sin dopar al lado tipo p de la juntura, también se genera una diferencia de potencial en esta juntura —de nuevo, la mitad de la caída de potencial en la juntura p - n —. En todas las juntas, las diferencias de potencial se ajustan de modo que no circule corriente por el circuito. Cualquier tipo de alambre que usen para conectar los dos lados de la juntura n - p , están produciendo dos nuevas juntas; mientras todas las juntas estén a la misma temperatura, los saltos de potencial en las juntas se compensan mutuamente, por lo que no habrá corriente circulando por el circuito. Sin embargo, sí ocurre —si tratan los detalles— que si alguna de las juntas está a temperatura diferente de las otras, circulará corriente. Unas juntas se calentarán y otras se enfriarán por el paso de la corriente, la energía térmica se convertirá en energía eléctrica. Este efecto es el responsable del funcionamiento de los termopares que se usan para medir temperaturas, y de los generadores termoeléctricos. También se usa el mismo efecto para hacer refrigeradores pequeños.

Si no podemos medir la diferencia de potencial entre los dos lados de una juntura n - p , ¿cómo podemos estar realmente seguros de que el gradiente de potencial que muestra la figura 14-9 en verdad existe? Una manera es iluminar la juntura. Cuando los fotones de la luz son absorbidos, éstos pueden producir un par electrón-hueco. En el fuerte campo eléctrico que hay en la juntura (igual a la pendiente de la curva de potencial de la figura 14-9) el hueco será dirigido a la región tipo p y el electrón a la región tipo n . Si los dos lados de la juntura están conectados a un circuito externo, estas cargas adicionales producirán una corriente. La energía de la luz se convertirá en energía eléctrica en la juntura. Las celdas solares, que generan potencia eléctrica para el funcionamiento de algunos de nuestros satélites, se basan en este principio.

En nuestra discusión del funcionamiento de una juntura de semiconductores, hemos supuesto que los huecos y los electrones actúan, más o menos, de manera independiente —excepto que, de algún modo, llegan al equilibrio estadístico—. Cuando describimos la corriente que se producía iluminando la juntura, supusimos que un electrón o un hueco, producidos en la región de la juntura, pasarían al cuerpo principal del cristal antes de aniquilarse con un portador de polaridad opuesta. En la vecindad inmediata de la juntura, donde la densidad de portadores de ambos signos es aproximadamente igual, el efecto de la aniquilación electrón-hueco (o, como se la llama frecuentemente, “recombinación”) es un efecto importante, y en un análisis detallado de juntas de semiconductores se debe tomar en cuenta adecuadamente. Hemos estado suponiendo que un hueco o un electrón producidos en una juntura tienen muchas probabilidades para pasar al cuerpo del cristal antes de recombinarse. El tiempo típico que tarda un electrón o un hueco en encontrar un compañero opuesto y aniquilarse está entre 10^{-3} y 10^{-7} segundos para materiales semiconductores típicos. Además, este tiempo es mucho mayor que el tiempo libre medio τ entre colisiones con los centros de dispersión en el cristal, que usamos en el análisis de la conductividad. En una

juntura n - p típica, el tiempo que un electrón o un hueco, formados en la región de la juntura, tardan en ser barridos hacia el cuerpo del cristal es, generalmente, mucho más corto que el tiempo de recombinación. En consecuencia, la mayoría de los pares contribuirán a una corriente externa.

14-5 Rectificación en una juntura de semiconductores

Queremos mostrar ahora cómo es que una juntura p - n puede funcionar como un rectificador. Si aplicamos un voltaje a la juntura, circulará una corriente mayor si la polaridad está en una dirección, pero la corriente será menor si se aplica el mismo voltaje en sentido opuesto. Si se aplica un voltaje alterno a la juntura, una corriente neta circula en una dirección —la corriente es “rectificada”—. Examinemos de nuevo lo que pasa en la condición de equilibrio descrita por las gráficas en la figura 14-9. En el material tipo p , hay una gran concentración N_p de portadores positivos. Estos portadores se difunden por todas partes, y cierto número de ellos se aproxima a la juntura cada segundo. Esta corriente de portadores positivos que se acerca a la juntura es proporcional a N_p . Sin embargo, la mayoría de ellos son regresados debido al alto escalón de potencial en la juntura, y sólo la fracción $e^{-qV/\kappa T}$ logra pasar. También hay una corriente de portadores positivos que se acerca a la juntura desde el otro lado. Esta corriente también es proporcional a la densidad de portadores positivos de la región tipo n , pero aquí la densidad de portadores es mucho menor que en el lado tipo p . Cuando los portadores positivos se aproximan a la juntura desde el lado tipo n , encuentran un escalón con pendiente negativa, e inmediatamente se deslizan hacia el lado tipo p de la juntura. Llamemos a esta corriente I_0 . En el equilibrio, las corrientes en ambas direcciones son iguales. Esperamos, entonces, la siguiente relación:

$$I_0 \propto N_p(\text{lado } n) = N_p(\text{lado } p)e^{-qV/\kappa T}. \quad (14.12)$$

Notarán que esta ecuación es, en realidad, la misma que la ecuación (14.10). Acabamos de deducirla de un modo diferente.

Supongan, sin embargo, que bajamos el voltaje en el lado n de la juntura en una cantidad ΔV —lo cual podemos hacer aplicando una diferencia de potencial externa a la juntura—. Ahora la diferencia de potencial a lo largo del escalón de potencial ya no es V , sino $V - \Delta V$. La corriente de portadores positivos desde el lado p hacia el lado n tendrá ahora esta diferencia de potencial en su factor exponencial. Llamando a esta corriente I_1 , tenemos

$$I_1 \propto N_p(\text{lado } p)e^{-q(V-\Delta V)/\kappa T}.$$

Esta corriente es mayor a I_0 , justamente por el factor $e^{q\Delta V/\kappa T}$. Por lo que tenemos la siguiente relación entre I_1 e I_0 :

$$I_1 = I_0 e^{+q\Delta V/\kappa T}. \quad (14.13)$$

La corriente desde el lado p aumenta exponencialmente con el voltaje ΔV aplicado externamente. Sin embargo, la corriente de portadores positivos del lado n permanece constante, siempre que ΔV no sea demasiado grande. Cuando se aproximen a la barrera, estos portadores encontrarán todavía una bajada de potencial y caerán al lado p . (Si ΔV fuese mayor que la diferencia de potencial natural V , las cosas cambiarían, pero no consideraremos lo que ocurriría a esos voltajes tan altos.) La corriente neta I de portadores positivos que circula a través de la juntura es, desde luego, la diferencia entre las corrientes de ambos lados:

$$I = I_0(e^{+q\Delta V/\kappa T} - 1). \quad (14.14)$$

La corriente neta I de huecos circula hacia la región tipo n . Ahí los huecos se difunden hacia el cuerpo de la región n , donde eventualmente se aniquilan debido a la mayoría de portadores tipo n —los electrones—. Los electrones que se pierden en esta aniquilación, se compensan por una corriente de electrones de la terminal externa del material tipo n .

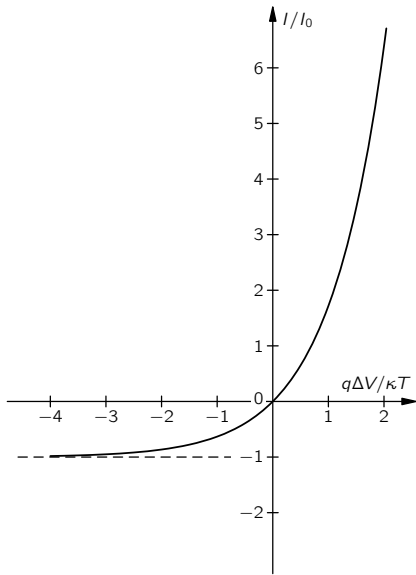


Figura 14-10. La corriente a través de la juntura en función del voltaje aplicado a ella.

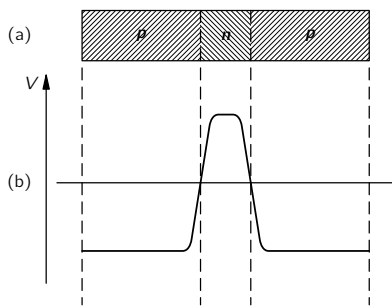


Figura 14-11. La distribución de potencial en un transistor sin voltajes aplicados.

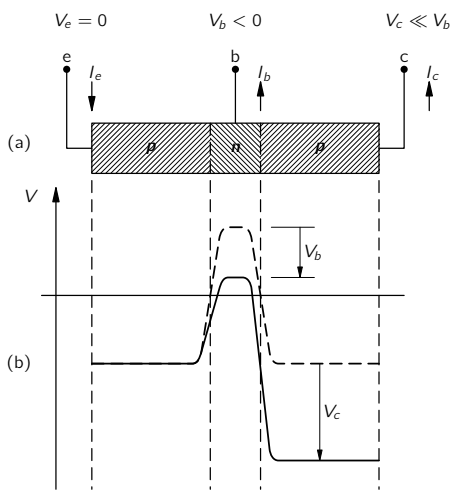


Figura 14-12. La distribución de potencial en un transistor en funcionamiento.

Cuando ΔV es cero, la corriente neta de la ecuación (14.14) es cero. Para ΔV positiva, la corriente aumenta rápidamente con el voltaje aplicado. Para ΔV negativa, la corriente invierte su signo, pero el término exponencial se hace rápidamente despreciable, y la corriente negativa nunca sobrepasa I_0 —que, con nuestras suposiciones, es bastante pequeña—. Esta contracorriente I_0 está limitada por la pequeña densidad de la minoría de los portadores tipo p en el lado n de la juntura.

Si hacen el mismo análisis para la corriente de portadores negativos que circula a través de la juntura, primero sin diferencia de potencial, y luego con una pequeña diferencia de potencial ΔV aplicada externamente, también obtienen una ecuación parecida a la (14.14) para la corriente neta de electrones. Como la corriente total es la suma de las corrientes debidas a los dos portadores, la ecuación (14.14) también sirve para la corriente total, siempre que identifiquemos a I_0 como la corriente máxima que puede circular con un voltaje invertido.

La relación voltaje-corriente característica de la ecuación (14.14) se muestra en la figura 14-10. Presenta el comportamiento típico de los diodos de estado sólido —como los que se usan en las computadoras modernas—. Debemos señalar que la ecuación (14.14) sólo es válida para voltajes pequeños. Para voltajes semejantes o mayores que la diferencia de voltaje interno natural V , entran en juego otros efectos, y la corriente ya no sigue una ecuación tan simple.

A propósito, ustedes recordarán que obtuvimos exactamente la misma ecuación (14.14), cuando estudiamos el “rectificador mecánico” —la rueda dentada con trinquete— en el capítulo 46 del volumen I. En las dos situaciones obtenemos las mismas ecuaciones, porque los procesos físicos básicos son bastante similares.

14-6 El transistor

Tal vez la aplicación más importante de los semiconductores sea el transistor. El transistor consiste en dos junturas de semiconductores muy cercanas. Su funcionamiento se basa, en parte, en los mismos principios que acabamos de discutir para el diodo semiconductor —la juntura rectificadora—. Supongan que hacemos una pequeña barra de germanio con tres regiones distintas, una región tipo p , una región tipo n y otra región tipo p , como muestra la figura 14-11(a). Esta combinación se denomina transistor p - n - p . Cada una de las dos junturas del transistor se comportará de modo muy parecido a lo descrito en la sección anterior. En particular, habrá un gradiente de potencial en cada juntura, que tendrá cierta caída de la región tipo n a cada región tipo p . Si las dos regiones tipo p tienen las mismas propiedades internas, la variación de potencial a lo largo del cristal será como lo muestra la gráfica de la figura 14-11(b).

Ahora imaginemos que conectamos cada una de las tres regiones a fuentes externas de voltaje, como lo muestra la parte (a) de la figura 14-12. Referiremos todos los voltajes a la terminal conectada a la región p de la izquierda, que tendrá, por definición, potencial cero. Llamaremos a esta terminal el *emisor*. La región tipo n es llamada la *base*, y está conectada a un potencial ligeramente negativo. La región tipo p de la derecha es llamada el *colector*, y está conectada a un potencial negativo algo mayor. En estas circunstancias, la variación de potencial a lo largo del cristal será como lo muestra la gráfica de la figura 14-12(b).

Veamos primero qué ocurre con los portadores positivos, ya que su comportamiento es el que determina principalmente el funcionamiento del transistor p - n - p . Como el emisor está a un potencial relativamente más positivo que la base, entonces circulará una corriente de portadores positivos de la región del emisor a la región de la base. Una corriente relativamente grande circula, puesto que tenemos una juntura funcionando con un “voltaje hacia adelante” —que corresponde a la mitad derecha de la gráfica en la figura 14-10—. En estas condiciones, los portadores positivos o huecos son “emitidos” de la región tipo p hacia la región tipo n . Podrían pensar que esta corriente saldría de la región tipo n por la terminal b de la base. No obstante, ahora viene el secreto del transistor. La región tipo n se hace muy delgada —típicamente 10^{-3} cm o menos, mucho más angosta que sus dimensiones transversales—. Esto significa que, a medida

que los huecos entran en la región tipo n , es muy probable que se difundan hasta la otra juntura antes de que se aniquilen con los electrones de la región tipo n . Cuando llegan a la frontera derecha de la región tipo n , encuentran un escalón de potencial hacia abajo, y caen inmediatamente en la región tipo p de la derecha. Este lado del cristal se llama colector porque “recolecta” los huecos después de que se han difundido a través de la región tipo n . En un transistor típico, toda la corriente de huecos, a excepción de una fracción del total, que deja el emisor y entra en la base, se recoge en la región colectora, y sólo la fracción que queda contribuye a la corriente neta de la base. Por supuesto, la suma de las corrientes de base y de colector es igual a la corriente de emisor.

Imaginen ahora lo que pasa si variamos ligeramente el potencial V_b en el terminal de la base. Como estamos en una parte bastante inclinada de la curva en la figura 14-10, una pequeña variación del potencial V_b ocasionará un cambio bastante grande en la corriente de emisor I_e . Como el voltaje del colector V_c es mucho más negativo que el voltaje de la base, estas pequeñas variaciones de potencial no afectarán sensiblemente al pronunciado escalón de potencial entre la base y el colector. La mayoría de los portadores positivos emitidos hacia la región n seguirán siendo atrapados por el colector. Así, a medida que variamos el potencial del electrodo de la base, habrá una variación correspondiente en la corriente de colector I_c . El punto esencial es, sin embargo, que la corriente de base I_b siempre permanece como una pequeña fracción de la corriente de colector. El transistor es un amplificador; una pequeña corriente I_b introducida en el electrodo de la base da una corriente grande —unas 100 veces mayor— en el electrodo del colector.

¿Y qué pasa con los electrones —los portadores negativos que hemos despreciado hasta ahora—? Noten primero que no esperamos que alguna corriente significativa de electrones circule entre la base y el colector. Con el gran voltaje negativo en el colector, los electrones de la base tendrían que subir un escalón muy alto de energía potencial, y la probabilidad para que lo hagan es muy pequeña. Hay una corriente muy pequeña de electrones hacia el colector.

Por otra parte, los electrones de la base *pueden* pasar a la región del emisor. De hecho, podrían esperar que la corriente de electrones en esta dirección fuese comparable a la corriente de huecos del emisor a la base. Esa corriente de electrones no es útil, y, por el contrario, es contraproducente, porque aumenta la corriente total de la base necesaria para una corriente dada de huecos hacia el colector. El transistor está, por lo tanto, diseñado para minimizar la corriente de electrones hacia el emisor. La corriente de electrones es proporcional a $N_n(\text{base})$, la densidad de los portadores negativos en el material de la base, mientras que la corriente de huecos desde el emisor depende de $N_p(\text{emisor})$, la densidad de los portadores positivos en la región del emisor. Por medio de un dopaje relativamente bajo del material tipo n , se puede hacer que $N_n(\text{base})$ sea mucho menor que $N_p(\text{emisor})$. (La región muy delgada de la base también ayuda mucho, porque el arrastre de huecos que el colector hace en esta región aumenta significativamente la corriente promedio de huecos del emisor a la base, dejando la corriente de electrones constante.) El resultado neto es que la corriente de electrones a través de la juntura emisor-base se puede hacer mucho menor que la corriente de huecos, por lo cual los electrones no desempeñan ningún papel importante en el funcionamiento del transistor $p-n-p$. Las corrientes están determinadas principalmente por el movimiento de los huecos, y el transistor funciona como un amplificador, como dijimos antes.

También se puede hacer un transistor intercambiando los materiales tipo p y tipo n en la figura 14-11. Por lo tanto, tenemos lo que se llama un transistor $n-p-n$. En el transistor $n-p-n$, las corrientes principales son transportadas por los electrones, los cuales circulan del emisor a la base, y de ahí al colector. Evidentemente, todos los argumentos que hemos hecho para el transistor $p-n-p$, también se aplican al transistor $n-p-n$, si se toman los potenciales de los electrodos con signos opuestos.

La aproximación de partículas independientes

15-1 Ondas de espín

En el capítulo 13 tratamos la teoría para la propagación de un electrón o de cualquier otra “partícula”, como una excitación atómica, a través de una red cristalina. En el capítulo 14 aplicamos la teoría a los semiconductores. Pero cuando hablamos de situaciones en las que hay muchos electrones, no consideramos ninguna interacción entre ellos. Hacer esto es, naturalmente, una aproximación. En este capítulo, discutiremos más ampliamente la idea de despreciar la interacción entre electrones. También, aprovecharemos la oportunidad para mostrar algunas otras aplicaciones de la teoría de propagación de partículas. Puesto que, en general, continuaremos despreciando las interacciones entre partículas, realmente hay muy pocas cosas nuevas en este capítulo, excepto por las nuevas aplicaciones. No obstante, el primer ejemplo que vamos a considerar, es uno en que es posible expresar de manera muy precisa las ecuaciones correctas cuando hay más de una “partícula” presente. A partir de esas ecuaciones, podremos ver cómo se hace la aproximación para despreciar las interacciones. Aun así, no analizaremos el problema con mucho cuidado.

Como primer ejemplo, consideraremos una “onda de espín” en un cristal ferromagnético. Hemos discutido la teoría del ferromagnetismo en el capítulo 36 del volumen II. A temperatura cero, todos los espines de los electrones que contribuyen al magnetismo en un cristal ferromagnético son paralelos. Hay una energía de interacción entre los espines, que es mínima cuando todos los espines están hacia abajo. Sin embargo, a una temperatura distinta de cero, hay cierta probabilidad para que algunos de los espines se hayan dado vuelta. Calculamos la probabilidad de una manera aproximada en el capítulo 36. Esta vez describiremos la teoría cuántica —y así verán lo que tendrían que hacer si quisieran resolver el problema de manera más exacta—. (Aun haremos algunas idealizaciones suponiendo que los electrones se localizan en los átomos, y que los espines interactúan solamente con los espines vecinos.)

Consideremos un modelo en el que los electrones de cada átomo están apareados, excepto por uno, por lo que todos los efectos magnéticos provienen de un electrón con espín $\frac{1}{2}$ por átomo. Más aún, imaginemos que estos electrones se localizan en los sitios atómicos de la red. El modelo corresponde vagamente al níquel metálico.

Además supondremos que hay una interacción entre cualquier par adyacente de electrones con espín, lo que origina un término en la energía del sistema

$$E = - \sum_{i,j} K \sigma_i \cdot \sigma_j, \quad (15.1)$$

donde las σ representan los espines, y la suma se extiende a todos los pares adyacentes de electrones. Ya hemos discutido esta clase de energía de interacción cuando consideramos la separación hiperfina del hidrógeno que se debe a la interacción de los momentos magnéticos del electrón y del protón del átomo de hidrógeno. La expresamos como $A \sigma_e \cdot \sigma_p$. Ahora bien, para un par dado, digamos los electrones de los átomos 4 y 5, el hamiltoniano sería $-K \sigma_4 \cdot \sigma_5$. Tenemos un término parecido para cada par, y el hamiltoniano es (como esperarían para energías clásicas) la suma de estos términos para cada par que interactúa. Escribimos la energía con el factor $-K$, para que así una K positiva corresponda

15-1 Ondas de espín

15-2 Dos ondas de espín

15-3 Partículas independientes

15-4 La molécula de benceno

15-5 Más química orgánica

15-6 Otros usos para la aproximación

al ferromagnetismo —esto es, la energía más baja se obtiene cuando los espines adyacentes son paralelos—. En un cristal real, puede haber otros términos, que son las interacciones de los *siguientes* vecinos más cercanos, etc., pero no necesitamos considerar tales complicaciones a estas alturas.

Con el hamiltoniano de la ecuación (15.1) tenemos una descripción completa del ferromagnetismo —dentro de nuestra aproximación— y deberíamos obtener las propiedades de la magnetización. Aun deberíamos ser capaces de calcular las propiedades termodinámicas debidas a la magnetización. Si podemos encontrar todos los niveles de energía, es posible encontrar las propiedades del cristal a la temperatura T a partir del principio de que la probabilidad para encontrar a un sistema en un estado dado de energía E es proporcional a $e^{-E/\kappa T}$. Este problema nunca se ha resuelto por completo.

Mostraremos algunas de las dificultades usando un ejemplo sencillo, en el cual todos los átomos están alineados —una red unidimensional—. Pueden extender fácilmente las ideas a tres dimensiones. En cada posición atómica hay un electrón que tiene dos estados posibles, ya sea espín hacia arriba o espín hacia abajo, y el sistema completo se describe diciendo cómo están distribuidos todos los espines. Suponemos que el hamiltoniano del sistema es el operador de la energía de interacción. Interpretando los vectores de espín de la ecuación (15.1) como los operadores sigma —o las matrices sigma— escribimos para la red lineal

$$\hat{H} = \sum_n -\frac{A}{2} \hat{\sigma}_n \cdot \hat{\sigma}_{n+1}. \quad (15.2)$$

En esta ecuación hemos escrito la constante igual a $A/2$ por conveniencia (de este modo, algunas de las ecuaciones que siguen serán exactamente como las del capítulo 13).

Ahora bien, ¿cuál es el estado más bajo del sistema? El estado de menor energía es aquel en que todos los espines son paralelos —digamos, todos hacia arriba—.† Podemos escribir a este estado como $|\cdots + + + \cdots\rangle$, o como $|\text{fnd}\rangle$ por estado “fundamental” o de menor energía. Es fácil calcular la energía para este estado. Una manera es expresar todos los vectores sigma en términos de $\hat{\sigma}_x$, $\hat{\sigma}_y$ y $\hat{\sigma}_z$, y calcular cuidadosamente lo que cada término del hamiltoniano le hace al estado fundamental, y luego sumar los resultados. Sin embargo, podemos usar un atajo. Vimos en la sección 12-2, que $\hat{\sigma}_i \cdot \hat{\sigma}_j$ se podía escribir en función del operador de intercambio de espín de Pauli en la forma:

$$\hat{\sigma}_i \cdot \hat{\sigma}_j = (2\hat{P}_{ij}^{\text{int espín}} - 1), \quad (15.3)$$

donde el operador $\hat{P}_{ij}^{\text{int espín}}$ intercambia los espines de los electrones i -ésimo y j -ésimo. Con esta sustitución, el hamiltoniano se convierte en

$$\hat{H} = -A \sum_n (\hat{P}_{n,n+1}^{\text{int espín}} - \frac{1}{2}). \quad (15.4)$$

Ahora es fácil ver lo que sucede a estados diferentes. Por ejemplo, si i y j están ambos hacia arriba, intercambiar los espines lo deja todo como estaba, por lo que $\hat{P}_{ij}$ actuando sobre el estado nos regresa el mismo estado, y es equivalente a multiplicar por $+1$. La expresión $(\hat{P}_{ij} - \frac{1}{2})$ es exactamente igual a un medio. (Desde ahora en adelante, omitiremos el superíndice descriptivo de $\hat{P}$.)

Para el estado fundamental, todos los espines están hacia arriba; por ello, si intercambian un par determinado de espines obtendrán el estado original. El estado fundamental es un estado estacionario. Si le aplican el operador hamiltoniano, obtendrán de nuevo el mismo estado multiplicado por una suma de términos, $-(A/2)$ para cada par de espines. Esto es, la energía del sistema en el estado fundamental es $-A/2$ por átomo.

† El estado fundamental aquí está realmente “degenerado”; hay otros estados con la misma energía —por ejemplo, todos los espines hacia abajo, o todos en cualquier otra dirección—. El más ínfimo campo externo en la dirección z dará una energía diferente a todos estos estados, y el que nosotros hemos escogido será el verdadero estado fundamental.

Después, nos gustaría considerar las energías de alguno de los estados excitados. Será conveniente medir las energías respecto al estado fundamental —esto es, escoger al estado fundamental como nuestro cero de energía—. Podemos hacer esto sumando la energía $A/2$ a cada término del hamiltoniano. Esto cambia solamente el “ $\frac{1}{2}$ ” de la ecuación (15.4) a “1”. Nuestro nuevo hamiltoniano es

$$\hat{H} = -A \sum_n (\hat{P}_{n,n+1} - 1). \quad (15.5)$$

Con este hamiltoniano, la energía del estado más bajo es cero; el operador de intercambio de espín es equivalente a multiplicar por uno (para el estado fundamental) que se cancela con el “1” en cada término.

Para describir estados que no sean el estado fundamental, necesitaremos un conjunto conveniente de estados base. Una buena aproximación es agrupar los estados de acuerdo a si un electrón tiene el espín hacia abajo, o si dos electrones tienen espín hacia abajo, y así sucesivamente. Hay, naturalmente, muchos estados con un espín hacia abajo. El espín hacia abajo podría estar en el átomo “4”, o en el átomo “5”, o en el átomo “6”... Podemos, en realidad, escoger justamente estos estados como nuestros estados base. Podríamos escribirlos así: $|4\rangle$, $|5\rangle$, $|6\rangle$... Sin embargo, después será más conveniente si identificamos el “átomo diferente” —el que tiene el electrón con espín hacia abajo— por su coordenada x . Esto es, definiremos al estado $|x_5\rangle$ como aquel en que todos los electrones tienen espines hacia arriba, excepto el electrón del átomo que está en x_5 , el cual tiene un electrón con espín hacia abajo (véase la figura 15-1). En general, $|x_n\rangle$ es el estado con un espín hacia abajo localizado en la coordenada x_n del átomo n -ésimo.

¿Qué efecto tiene el hamiltoniano (15.5) sobre el estado $|x_5\rangle$? Uno de los términos del hamiltoniano es $-A(\hat{P}_{7,8} - 1)$, por ejemplo. El operador $\hat{P}_{7,8}$ intercambia los dos espines de los átomos adyacentes 7 y 8. Pero en el estado $|x_5\rangle$ ambos están hacia arriba, por lo que no ocurre nada; $\hat{P}_{7,8}$ es equivalente a multiplicar por 1:

$$\hat{P}_{7,8} |x_5\rangle = |x_5\rangle.$$

De donde

$$(\hat{P}_{7,8} - 1) |x_5\rangle = 0.$$

Así, todos los términos del hamiltoniano nos dan cero —excepto aquellos en los que interviene el átomo 5, por supuesto—. Para el estado $|x_5\rangle$, el operador $\hat{P}_{4,5}$ intercambia los espines de los átomos 4 (hacia arriba) y 5 (hacia abajo). El resultado es el estado con todos los espines hacia arriba, excepto el del átomo 4. Esto es

$$\hat{P}_{4,5} |x_5\rangle = |x_4\rangle.$$

Del mismo modo

$$\hat{P}_{5,6} |x_5\rangle = |x_6\rangle.$$

Entonces, los únicos términos del hamiltoniano que sobreviven son $-A(\hat{P}_{4,5} - 1)$ y $-A(\hat{P}_{5,6} - 1)$. Actuando sobre $|x_5\rangle$, producen $-A|x_4\rangle + A|x_5\rangle$ y $-A|x_6\rangle + A|x_5\rangle$, respectivamente. El resultado es

$$\hat{H} |x_5\rangle = -A \sum_n (\hat{P}_{n,n+1} - 1) |x_5\rangle = -A\{|x_6\rangle + |x_4\rangle - 2|x_5\rangle\}. \quad (15.6)$$

Cuando el hamiltoniano actúa sobre el estado $|x_5\rangle$ origina cierta amplitud para estar en los estados $|x_4\rangle$ y $|x_6\rangle$. Así, hay una cierta amplitud para que el espín hacia abajo salte al átomo vecino. Por lo tanto, a causa de la interacción entre los espines, si comenzamos con un espín hacia abajo, entonces existe cierta probabilidad para que después sea otro el que esté hacia abajo. Cuando el hamiltoniano opera sobre el estado general $|x_n\rangle$ nos da

$$\hat{H} |x_n\rangle = -A\{|x_{n+1}\rangle + |x_{n-1}\rangle - 2|x_n\rangle\}. \quad (15.7)$$

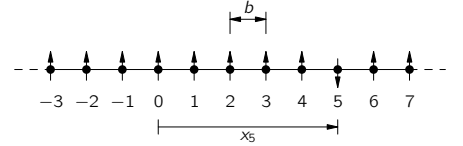


Figura 15-1. El estado base $|x_5\rangle$ de un arreglo lineal de espines. Todos los espines están hacia arriba, excepto el de la posición x_5 , el cual está hacia abajo.

Noten que, en particular, si tomamos un conjunto completo de estados base con sólo un espín hacia abajo, únicamente se combinarán entre ellos. El hamiltoniano nunca combinará estos estados con otros que tengan más espines hacia abajo. Por lo que, siempre que sólo intercambien espines, el número total de espines hacia abajo no cambiará.

Será conveniente usar la notación matricial para el hamiltoniano, digamos $H_{n,m} \equiv \langle x_n | \hat{H} | x_m \rangle$; la ecuación (15.7) es equivalente a

$$\begin{aligned} H_{n,n} &= 2A; \\ H_{n,n+1} &= H_{n,n-1} = -A; \\ H_{n,m} &= 0, \quad \text{para } |n-m| > 1. \end{aligned} \quad (15.8)$$

Ahora, ¿cuáles son los niveles de energía para los estados con un espín hacia abajo? Como siempre, decimos que C_n es la amplitud para que algún estado $|\psi\rangle$ esté en el estado $|x_n\rangle$. Si $|\psi\rangle$ va a ser un estado de energía definida, todas las C_n deben variar del mismo modo con el tiempo, o sea,

$$C_n = a_n e^{-iEt/\hbar}. \quad (15.9)$$

Podemos introducir esta solución de prueba en nuestra ecuación hamiltoniana

$$i\hbar \frac{dC_n}{dt} = \sum_m H_{nm} C_m, \quad (15.10)$$

usando la ecuación (15.8) para los elementos matriciales. Por supuesto, obtenemos un número infinito de ecuaciones, pero todas se pueden escribir en la forma

$$Ea_n = 2Aa_n - Aa_{n-1} - Aa_{n+1}. \quad (15.11)$$

Tenemos exactamente el mismo problema que resolvimos en el capítulo 13, excepto que donde teníamos E_0 ahora tenemos $2A$. Las soluciones corresponden a amplitudes C_n (amplitud para espín hacia abajo) que se propagan a lo largo de la red con una constante de propagación k y una energía

$$E = 2A(1 - \cos kb), \quad (15.12)$$

donde b es la constante de la red.

Las soluciones de energía definida corresponden a “ondas” de espín hacia abajo —llamadas “ondas de espín”—. Y para cada longitud de onda, hay una energía correspondiente. Para longitudes de onda grandes (k pequeñas) la energía varía como

$$E = Ab^2 k^2. \quad (15.13)$$

Exactamente como antes, podemos considerar un paquete de ondas localizado (que contiene, sin embargo, sólo longitudes de ondas largas) que corresponde a un electrón con espín hacia abajo, en un lugar de la red. Este espín hacia abajo se comportará como una “partícula”. Como su energía se relaciona con k mediante (15.13), la “partícula” tendrá una masa efectiva:

$$m_{\text{ef}} = \frac{\hbar^2}{2Ab^2}. \quad (15.14)$$

A estas “partículas” algunas veces se les llama “magnones”.

15-2 Dos ondas de espín

Ahora discutiremos lo que sucede si hay *dos* espines hacia abajo. De nuevo, escogemos un conjunto de estados base. Escogemos estados en los que hay espines hacia abajo en dos posiciones atómicas, como el estado que se muestra en la figura 15-2. Podemos identificar a este estado por las coordenadas x de los dos lugares con espines hacia abajo. El de la figura se puede llamar $|x_2, x_5\rangle$. En general,

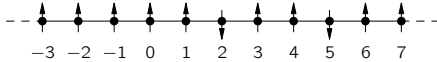


Figura 15-2. Un estado con dos espines hacia abajo.

los estados base son $|x_n, x_m\rangle$ —¡un conjunto doblemente infinito!—. En este sistema de descripción, el estado $|x_4, x_9\rangle$ y el estado $|x_9, x_4\rangle$ son exactamente el mismo estado, ya que cada uno indica simplemente que hay un espín hacia abajo en 4 y otro en 9; el orden no significa nada. Además, el estado $|x_4, x_4\rangle$ no tiene sentido, no existe tal cosa. Podemos describir cualquier estado $|\psi\rangle$ dando las amplitudes para que se encuentre en cada uno de los estados base. Entonces, $C_{m,n} = \langle x_m, x_n | \psi \rangle$ significa ahora la amplitud para que un sistema en el estado $|\psi\rangle$ esté en un estado en el que los átomos m -ésimo y n -ésimo tengan espín hacia abajo. Las complicaciones que surgen ahora no son complicaciones conceptuales —son simplemente complejidades de cómo llevar registros—. (Una de las complejidades de la mecánica cuántica es cómo llevar registros. Si tenemos más y más espines hacia abajo, la notación se hace más y más complicada, con una gran cantidad de índices y las ecuaciones siempre parecen muy aterradoras; pero las ideas no son necesariamente más complicadas que en el caso sencillo.)

Las ecuaciones de movimiento para el sistema de espín son las ecuaciones diferenciales para las $C_{n,m}$. Éstas son

$$i\hbar \frac{dC_{n,m}}{dt} = \sum_{i,j} (H_{nm,ij}) C_{ij}. \quad (15.15)$$

Supongan que queremos encontrar los estados estacionarios. Como de costumbre, las derivadas respecto al tiempo nos dan E por las amplitudes, y las $C_{m,n}$ se pueden remplazar por los coeficientes $a_{m,n}$. Después, tenemos que calcular cuidadosamente el efecto de H sobre un estado con espines m y n hacia abajo. No es difícil resolver. Supongan por un momento que m y n están lo suficientemente separados, para que no tengamos que preocuparnos de la dificultad evidente. La operación de intercambio en el lugar x_n trasladará el espín hacia abajo hasta el átomo $(n+1)$ o hasta el $(n-1)$, y, por lo tanto, hay una amplitud para que el estado actual proceda del estado $|x_m, x_{n+1}\rangle$ y, también, una amplitud para que proceda del estado $|x_m, x_{n-1}\rangle$. O podría haber sido el otro espín el que se trasladó; por lo que hay una cierta amplitud para que $C_{m,n}$ proceda de $C_{m+1,n}$ o de $C_{m-1,n}$. Estos efectos deberían ser todos iguales. El resultado final para la ecuación hamiltoniana sobre $C_{m,n}$ es

$$Ea_{m,n} = -A(a_{m+1,n} + a_{m-1,n} + a_{m,n+1} + a_{m,n-1}) + 4Aa_{m,n}. \quad (15.16)$$

Esta ecuación es correcta excepto en dos ocasiones. Si $m = n$ no tenemos ecuación, y si $m = n \pm 1$ faltarían dos términos en la ecuación (15.16). *Vamos a ignorar estas excepciones.* Simplemente ignoramos el hecho de que unas pocas de estas ecuaciones están levemente alteradas. Después de todo, se supone que el cristal es infinito, y tenemos un número infinito de términos; despreciar unos pocos no debe importar mucho. Así que, en una primera aproximación, olvidaremos las ecuaciones alteradas. En otras palabras, suponemos que la ecuación (15.16) es verdadera para cualesquiera m y n , aun para m y n consecutivos. *Ésta es la parte esencial de nuestra aproximación.*

De este modo, no es difícil encontrar la solución. Obtenemos inmediatamente

$$C_{m,n} = a_{m,n} e^{-iEt/\hbar}, \quad (15.17)$$

con

$$a_{m,n} = (\text{const}) e^{ik_1 x_m} e^{ik_2 x_n}, \quad (15.18)$$

donde

$$E = 4A - 2A \cos k_1 b - 2A \cos k_2 b. \quad (15.19)$$

Piensen por un momento lo que ocurriría si tuviésemos dos ondas de espín *solas e independientes* (como en la sección anterior) correspondientes a $k = k_1$ y $k = k_2$; tendrían las energías, según la ecuación (15.12),

$$\epsilon_1 = (2A - 2A \cos k_1 b)$$

y

$$\epsilon_2 = (2A - 2A \cos k_2 b).$$

Noten que la energía E en la ecuación (15.19) es sólo la suma de ambas

$$E = \epsilon(k_1) + \epsilon(k_2). \quad (15.20)$$

En otras palabras, podemos considerar nuestra solución en esta forma. Hay dos partículas —esto es, dos ondas de espín—. Una de ellas tiene un momentum descrito por k_1 , la otra por k_2 , y la energía del sistema es la suma de la energía de los dos objetos. Las dos partículas actúan independientemente por completo. Esto es todo lo que ocurre.

Hemos hecho algunas aproximaciones, naturalmente, pero no deseamos discutir la precisión de nuestra respuesta en este momento. Sin embargo, pueden imaginarse que en un cristal de tamaño razonable con miles de millones de átomos —y, por lo tanto, con miles de millones de términos en el hamiltoniano— el desprestigiar algunos términos no producirá un error apreciable. Si tuviésemos una cantidad suficiente de espines hacia abajo de modo que hubiera una densidad apreciable, entonces sí tendríamos que preocuparnos por las correcciones.

[Es muy interesante, se puede dar una solución exacta si hay solamente *dos* espines hacia abajo. El resultado no es particularmente importante. Pero es interesante que las ecuaciones se puedan resolver exactamente para este caso. La solución es:

$$a_{m,n} = \exp[ik_c(x_m + x_n)] \sin(k|x_m - x_n|), \quad (15.21)$$

con la energía

$$E = 4A - 2A \cos k_1 b - 2A \cos k_2 b,$$

y con los números de onda k_c y k relacionados con k_1 y k_2 mediante

$$k_1 = k_c - k, \quad k_2 = k_c + k. \quad (15.22)$$

Esta solución incluye la “interacción” entre los dos espines. Describe el hecho de que cuando los espines se acercan, hay una cierta probabilidad para la dispersión. Los espines se comportan de un modo muy parecido a las partículas con interacciones. Pero la teoría detallada de su dispersión rebasa los límites de lo que queremos decir aquí.]

15-3 Partículas independientes

En la sección anterior escribimos un hamiltoniano, ecuación (15.15), para un sistema de dos partículas. Después, usando una aproximación que es equivalente a desprestigiar cualquier “interacción” entre las dos partículas, encontramos los estados estacionarios descritos por las ecuaciones (15.17) y (15.18). Este estado es simplemente el producto de dos estados de partículas solas. Sin embargo, la solución que hemos dado en la ecuación (15.18) para $a_{m,n}$ no es realmente satisfactoria. Hemos señalado antes muy específicamente que el estado $|x_9, x_4\rangle$ *no* es diferente del estado $|x_4, x_9\rangle$ —el *orden* de x_m y x_n no tiene importancia—. En general, la expresión algebraica para la amplitud $C_{m,n}$ no debe cambiar si intercambiamos los valores de x_m y x_n , puesto que no cambia el estado. De cualquier modo, debería representar la amplitud para encontrar un espín hacia abajo en x_m y un espín hacia abajo en x_n . Pero noten que (15.18) *no* es simétrica respecto a x_m y x_n —ya que k_1 y k_2 pueden ser diferentes en general—.

El problema es que no hemos forzado a la solución de la ecuación (15.15) para que satisfaga esta condición adicional. Afortunadamente es fácil corregirlo. En primer lugar, noten que una solución a la ecuación hamiltoniana tan eficaz como (15.18) es

$$a_{m,n} = K e^{ik_2 x_m} e^{ik_1 x_n}. \quad (15.23)$$

Incluso tiene la misma energía que obtuvimos para (15.18). Cualquier combinación lineal de (15.18) y (15.23) es también una solución, y además tiene

la energía dada por (15.19). La solución que deberíamos haber escogido —a causa de nuestro requisito de simetría— es justamente la suma de (15.18) y (15.23):

$$a_{m,n} = K[e^{ik_1x_m}e^{ik_2x_n} + e^{ik_2x_m}e^{ik_1x_n}]. \quad (15.24)$$

En esta forma, dados k_1 y k_2 cualesquiera, la amplitud $C_{m,n}$ es independiente del modo en que coloquemos x_m y x_n —si se nos ocurriera definir a x_m y x_n en orden inverso, obtendríamos la misma amplitud—. Nuestra interpretación de la ecuación (15.24) en términos de “magnones” también debe ser diferente. Ya no podemos decir que la ecuación representa *una* partícula con número de onda k_1 y una *segunda* partícula con número de onda k_2 . La amplitud (15.24) representa *un* único estado con dos partículas (magnones). El *estado* es caracterizado por los dos números de onda k_1 y k_2 . Nuestra solución se parece a un estado compuesto de una partícula con momentum $p_1 = \hbar k_1$ y otra partícula con momentum $p_2 = \hbar k_2$, pero en nuestro estado no podemos decir cuál partícula es cuál.

Ahora, esta discusión les debe recordar al capítulo 4 y nuestra historia de las partículas idénticas. Les hemos estado mostrando que las partículas de las ondas de espín —los magnones— se comportan en forma parecida a bosones idénticos. Todas las amplitudes deben ser simétricas en las coordenadas de las dos partículas —que es como decir que si “intercambiamos las dos partículas”, volvemos a obtener la misma amplitud y con el mismo signo—. Pero, pueden estar pensando, ¿por qué escogimos *sumar* los dos términos para formar la ecuación (15.24)? ¿Por qué no escogimos *restar*? Con un signo menos, intercambiar x_m y x_n solamente cambiaría el signo de $a_{m,n}$, lo cual no importa. Pero intercambiar x_m y x_n *no altera nada* —todos los electrones del cristal están exactamente donde estaban antes, y por ello no hay ninguna razón para que cambie el signo de la amplitud—. Los magnones se comportan como bosones.†

Los puntos principales de esta discusión han sido dos: primero, mostrarles algo acerca de las ondas de espín, y segundo, mostrarles un estado cuya amplitud es el *producto* de dos amplitudes, y cuya energía es la *suma* de las energías correspondientes a las dos amplitudes. Para *partículas independientes*, la amplitud es el producto y la energía es la suma. Pueden comprobar fácilmente por qué la energía es la suma. La energía es el coeficiente de t en un exponencial imaginario —es proporcional a la frecuencia—. Si dos objetos están haciendo algo, uno de ellos con la amplitud $e^{-iE_1t/\hbar}$ y otro con la amplitud $e^{-iE_2t/\hbar}$, y si la amplitud para que ambos sucedan juntos es el producto de las amplitudes para cada uno de ellos, entonces hay una única frecuencia en el producto que es la suma de las dos frecuencias. La energía que corresponde al producto de las amplitudes es la suma de las dos energías.

Nos hemos extendido demasiado para decirles algo muy sencillo. Cuando no toman en cuenta ninguna interacción entre las partículas, pueden considerar cada una de ellas independientemente. Pueden existir individualmente en los diversos estados diferentes que tendrían solas, y cada una contribuirá con la energía que habría tenido si estuviera sola. Sin embargo, deben recordar que, si son partículas idénticas, se pueden comportar como bosones o como fermiones, según sea el problema. Los dos electrones adicionales añadidos al cristal, por ejemplo, tendrían que comportarse como fermiones. Cuando se intercambian las posiciones de dos electrones, la amplitud debe cambiar de signo. En la ecuación correspondiente a (15.24) debería haber un signo menos entre los dos términos del lado derecho. Como consecuencia, dos fermiones no pueden estar exactamente en el mismo estado —con espines iguales y k iguales—. La amplitud para este estado es cero.

15-4 La molécula de benceno

Aunque la mecánica cuántica provee las leyes básicas que determinan la estructura de las moléculas, estas leyes sólo se pueden aplicar de manera exacta a los

† En general, las cuasipartículas del tipo que estamos discutiendo pueden actuar como bosones o fermiones, y como para las partículas libres, las que tienen espín entero son bosones y las que lo tienen racional son fermiones. El “magnón” indica un electrón con espín hacia arriba que se voltea. El *cambio* en el espín es *uno*. El magnón tiene espín entero y es un bosón.

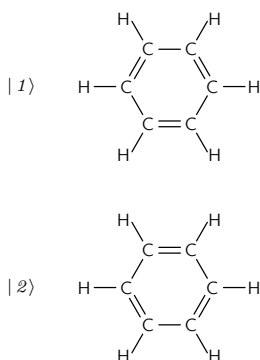


Figura 15-3. Los dos estados base para la molécula de benceno usados en el capítulo 10.

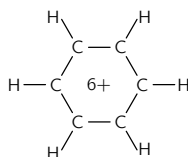


Figura 15-4. Un anillo de benceno con seis electrones removidos.

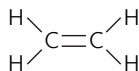


Figura 15-5. La molécula de etileno.

compuestos más simples. Por tanto, los químicos han elaborado varios métodos aproximados para calcular algunas propiedades de moléculas complicadas. Quereamos mostrarles ahora cómo se usa la aproximación de partículas independientes en la química orgánica. Comenzaremos con la molécula de benceno.

Discutimos a la molécula de benceno desde otro punto de vista en el capítulo 10. Ahí hicimos una idea aproximada de la molécula como sistema de dos estados, con los estados base que muestra la figura 15-3. Hay un anillo de seis carbonos, cada uno de ellos con un hidrógeno enlazado. Con la idea convencional de enlaces de valencia, es necesario suponer enlaces dobles entre la mitad de los átomos de carbono, y en el estado de energía más baja encontramos las dos posibilidades mostradas en la figura. Hay también otros estados de energía más alta. Cuando discutimos el benceno en el capítulo 10, sólo consideramos a los dos estados y olvidamos los demás. Encontramos que la energía del estado fundamental de la molécula no era la energía de uno de los estados de la figura, sino que era más baja en una cantidad proporcional a la amplitud para cambiar de uno de estos estados al otro.

Ahora consideraremos la misma molécula, pero desde un punto de vista completamente diferente —usando una aproximación distinta—. Los dos puntos de vista nos darán respuestas diferentes, pero si mejoramos cualquiera de las aproximaciones, ésta nos conducirá a la verdad, una descripción válida para el benceno. No obstante, si no nos preocupamos por mejorarlas, lo cual ocurre por lo general, entonces no deben sorprenderse si las dos descripciones no concuerdan de manera exacta. Por lo menos, mostraremos que también con el nuevo punto de vista, se tiene que la energía más baja de la molécula de benceno es menor que cualquiera de las estructuras con tres enlaces dobles de la figura 15-3.

Nos interesa usar ahora la siguiente descripción. Supongan que imaginamos a los seis átomos de carbono de la molécula de benceno con enlaces simples como en la figura 15-4. Hemos quitado seis electrones —ya que cada enlace equivale a un par de electrones—; por ello, tenemos una molécula de benceno seis veces ionizada. Ahora consideraremos lo que sucede cuando volvemos a colocar los seis electrones, uno cada vez, imaginando que cada uno puede moverse libremente alrededor del anillo. Suponemos también que todos los enlaces de la figura 15-4 están saturados, y no tenemos que considerarlos en adelante.

¿Qué sucede cuando volvemos a colocar un electrón en el ion molecular? Podría, naturalmente, colocarse en cualquiera de las seis posiciones alrededor del anillo —correspondientes a seis estados base—. También tendría una cierta amplitud, A por ejemplo, para ir de una posición a la próxima. Si analizamos los estados estacionarios, habría ciertos niveles de energía posibles. Esto es solamente con un electrón.

Luego colocamos un segundo electrón, y hacemos la aproximación más ridícula que se puedan imaginar —lo que un electrón hace no se ve afectado por lo que el otro está haciendo—. Es natural que ellos realmente interactúen; se repelen debido a la fuerza de Coulomb, y más aún, cuando se encuentren ambos en el mismo lugar, tendrán una energía considerablemente diferente al doble de la energía si sólo uno estuviese ahí. Ciertamente, la aproximación de partículas independientes no es apropiada cuando sólo hay seis lugares —en particular, cuando queremos colocar *seis* electrones—. No obstante, los químicos orgánicos han aprendido mucho al hacer esta clase de aproximaciones.

Antes de tratar a la molécula de benceno en detalle, consideremos un ejemplo más sencillo —la molécula de etileno, que tiene solamente dos átomos de carbono con dos átomos de hidrógeno a cada lado, como muestra la figura 15-5—. Esta molécula tiene un enlace “adicional”, que implica dos electrones, entre los átomos de carbono. Ahora quitamos uno de estos electrones; ¿qué tenemos? Podemos considerarlo como un sistema de dos estados —el electrón que queda puede estar en un carbono o el otro—. Podemos analizarlo como un sistema de dos estados. Las energías posibles para el único electrón son $(E_0 - A)$ o bien $(E_0 + A)$, como muestra la figura 15-6.

Ahora añadimos el segundo electrón. Bien, si tenemos dos electrones, entonces podemos colocar al primero en el estado inferior y el segundo en el superior.

Pero hemos olvidado algo. Cada uno de los estados es en verdad doble. Cuando decimos que hay un estado posible con energía $(E_0 - A)$, hay dos realmente. Dos electrones pueden ir al mismo estado, si uno de ellos tiene su espín hacia arriba y el otro hacia abajo. (No podemos colocar más a causa del principio de exclusión.) Así, realmente hay dos estados posibles de energía $(E_0 - A)$. Podemos hacer un diagrama, como en la figura 15-7, que indique los dos niveles de energía y su grado de ocupación. En el estado de menor energía, ambos electrones estarán en el estado inferior con sus espines opuestos. La energía del enlace adicional de la molécula de etileno es, por lo tanto, $2(E_0 - A)$ si despreciamos la interacción entre los dos electrones.

Volvamos al benceno. Cada uno de los dos estados en la figura 15-3 tiene tres enlaces dobles. Cada uno de ellos es exactamente como el enlace del etileno, y contribuye con $2(E_0 - A)$ a la energía, si E_0 es ahora la energía de colocar un electrón en una de las posiciones del benceno, y A es la amplitud para cambiarlo a la posición vecina. Entonces, la energía sería aproximadamente $6(E_0 - A)$. Pero cuando estudiamos al benceno antes, obtuvimos que la energía era menor que la energía de la estructura con tres enlaces adicionales. Veamos si la energía del benceno resulta menor que los tres enlaces desde nuestro nuevo punto de vista.

Comenzamos con el anillo de benceno seis veces ionizado y añadimos un electrón. Tenemos así un sistema de seis estados. Todavía no hemos resuelto un sistema como éste, pero sabemos qué hacer. Podemos escribir seis ecuaciones para las seis amplitudes, y así sucesivamente. Pero ahorrémonos algo de trabajo—considerando que ya hemos resuelto el problema, cuando resolvimos el problema de un electrón en una línea infinita de átomos—. Naturalmente, el benceno no es una línea infinita, tiene 6 posiciones atómicas en círculo. Pero imaginen que abrimos el círculo y formamos una línea, y numeramos a los átomos del 1 hasta 6. En una línea infinita, la posición siguiente sería 7, pero si ésta la hacemos idéntica a la número 1 y así sucesivamente, la situación sería exactamente como la del anillo de benceno. En otras palabras, podemos tomar la solución para una línea infinita *con un requisito añadido*, que la solución debe ser periódica con un ciclo de longitud igual a seis átomos. Según el capítulo 13, el electrón en una línea tiene estados de energía definida cuando la amplitud en cada posición es $e^{ikx_n} = e^{ikbn}$. Para cada k la energía es

$$E = E_0 - 2A \cos kb. \quad (15.25)$$

Ahora queremos usar solamente aquellas soluciones que se repiten cada 6 átomos. Hagamos primero el caso general para un anillo con N átomos. Si la solución debe tener un periodo de espaciamiento atómico N , e^{ikbN} debe ser igual a uno; o sea, kbN debe ser múltiplo de 2π . Si decimos que s representa cualquier entero, nuestra condición es que

$$kbN = 2\pi s. \quad (15.26)$$

Hemos visto antes que no tiene ningún sentido tomar k fuera del intervalo $\pm\pi/b$. Esto significa que obtenemos todos los estados posibles al tomar valores de s en el intervalo $\pm N/2$.

Encontramos que para un anillo con N átomos, hay N estados de energía definidos† con números de onda k_s dados por

$$k_s = \frac{2\pi}{Nb} s. \quad (15.27)$$

Cada estado tiene la energía (15.25). Tenemos una línea espectral de posibles niveles de energía. El espectro para el benceno ($N = 6$) se muestra en la figura 15-8(b). (Los números entre paréntesis indican el número de estados *diferentes* con la misma energía.)

Hay una forma excelente para visualizar los seis niveles de energía, como hemos mostrado en la parte (a) de la figura. Imaginen un círculo centrado en

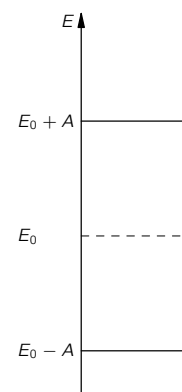


Figura 15-6. Los niveles de energía posibles para los electrones “adicionales” en la molécula de etileno.

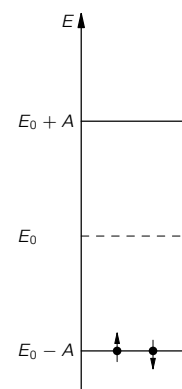


Figura 15-7. En el enlace adicional de molécula de etileno, dos electrones (uno con espín hacia arriba y otro con espín hacia abajo) pueden ocupar el nivel de energía más baja.

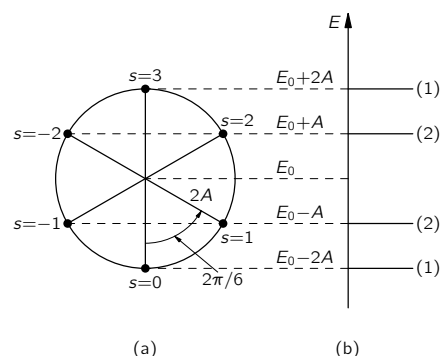


Figura 15-8. Los niveles de energía en un anillo con seis posiciones electrónicas (por ejemplo, un anillo de benceno).

† Podrían pensar que para un número par N hay $N + 1$ estados. Esto no es así, ya que $s = \pm N/2$ da el mismo estado.

el nivel E_0 y de radio $2A$. Si empezamos abajo y marcamos seis arcos iguales (con ángulos desde el punto inferior de $k_s b = 2\pi s/N$, o $2\pi s/6$ para el benceno), entonces las alturas verticales de los puntos en el círculo son las soluciones de la ecuación (15.25). Estos seis puntos representan a los seis estados posibles. El nivel de energía menor es $(E_0 - 2A)$; hay dos estados con la misma energía $(E_0 - A)$, y así sucesivamente†. Éstos son estados posibles para un electrón. Si tenemos más de un electrón, entonces dos —con espines opuestos— pueden estar en cada estado.

Para la molécula de benceno tenemos que colocar seis electrones. Para el estado fundamental, éstos irán a los estados de energía más baja posible —dos para $s = 0$, dos para $s = +1$ y dos para $s = -1$ —. Según la aproximación de partículas independientes, la energía del estado fundamental es

$$\begin{aligned} E_{\text{fnd}} &= 2(E_0 - 2A) + 4(E_0 - A) \\ &= 6E_0 - 8A. \end{aligned} \tag{15.28}$$

La energía es verdaderamente menor que la de los tres enlaces dobles separados —en la cantidad $2A$ —.

Al comparar la energía del benceno con la energía del etileno, es posible determinar A . Resulta ser de 0.8 electronvoltios, o, en las unidades que les gustan a los químicos, 18 kilocalorías por mol.

Podemos usar esta descripción para calcular, o entender, otras propiedades del benceno. Por ejemplo, con la figura 15-8, podemos discutir la excitación del benceno debida a la luz. ¿Qué sucedería si tratamos de excitar a uno de los electrones? Podría trasladarse a uno de los estados más altos vacíos. La energía más baja de excitación sería una transición del nivel completo más alto, al nivel vacío más bajo. Esto requiere la energía $2A$. El benceno absorberá luz de frecuencia ν cuando $h\nu = 2A$. También habrá absorción de fotones con energías $3A$ y $4A$. No hay necesidad de decir que se ha medido el espectro de absorción del benceno, y que el diagrama de líneas espectrales es más o menos correcto, excepto en que la transición más baja ocurre en el ultravioleta, y para ajustarse a los datos, se tendría que escoger un valor de A entre 1.4 y 2.4 electronvoltios. Esto es, el valor numérico de A es dos o tres veces mayor de lo que predice la energía de enlace químico.

Lo que el químico hace en situaciones como ésta, es analizar muchas moléculas similares y obtener algunas reglas empíricas. Así aprende, por ejemplo: para calcular la energía de enlace se debe usar tal y tal valor para A , pero para obtener el espectro de absorción aproximadamente bien, se debe usar otro valor de A . Esto puede parecerles un poco absurdo. No es muy satisfactorio desde el punto de vista del físico que trata de entender a la naturaleza por medio de principios básicos. Pero el problema de un químico es diferente. Él debe tratar de adivinar con anterioridad lo que va a suceder con moléculas que aún no se han formado, o que no se las comprende completamente. Por ello necesita una serie de reglas empíricas, sin importarle de donde provengan. Así, vemos que él usa la teoría de un modo bastante diferente al físico. Toma ecuaciones que poseen alguna sombra de verdad, pero tiene que alterar las constantes que aparecen —haciendo correcciones empíricas—.

En el caso del benceno, la principal razón para la inconsistencia, es nuestra suposición de que los electrones son independientes —la teoría con que empezamos no es realmente adecuada—. No obstante, tiene alguna sombra de verdad porque parece que sus resultados van en la dirección correcta. Con esas ecuaciones más algunas reglas empíricas —que incluyen varias excepciones— el químico orgánico atraviesa el lodazal tan complicado que ha escogido para estudiar. (No olviden que la razón por la que el físico pueda realmente calcular a partir de principios básicos, es que solamente escoge problemas sencillos. Nunca resuelve un problema con 42 electrones y ni siquiera uno con sólo 6. Hasta ahora, solamente

† Cuando hay dos estados (los cuales tendrán diferentes distribuciones de amplitud) con la misma energía, decimos que los dos estados están “degenerados”. Noten que *cuatro* electrones pueden tener la energía $E_0 - A$.

ha sido capaz de calcular el átomo de hidrógeno y el átomo de helio de un modo razonablemente exacto.)

15-5 Más química orgánica

Veamos cómo podemos utilizar las mismas ideas para estudiar otras moléculas. Consideren, por ejemplo, una molécula de 1,3-butadieno —se ha dibujado en la figura 15-9 de acuerdo a la imagen común del enlace de valencia—.

Podemos hacer el mismo juego con los cuatro electrones adicionales, que corresponden a los dos enlaces dobles. Si quitamos cuatro electrones, tenemos cuatro átomos de carbono en una línea. Ya saben cómo se resuelve una línea. Ustedes dirán, “¡Oh no! Sólo sé resolver una línea *infinita*”. Pero las soluciones de una línea infinita también incluyen las de una línea finita. Observen. Sea N el número de átomos en la línea y numerémoslos de 1 a N , como lo muestra la figura 15-10. Al escribir las ecuaciones para la amplitud en la posición 1, no tendrán un término proveniente de la posición 0. Del mismo modo, la ecuación para la posición N diferiría de la que usamos para una línea infinita, porque no habría nada proveniente de la posición $N + 1$. Pero supongan que podemos obtener una solución para la línea infinita que tenga la siguiente propiedad: la amplitud para estar en el átomo 0 es cero y la amplitud para estar en el átomo $(N + 1)$ es también cero. Así, el conjunto de ecuaciones para todas las posiciones desde 1 hasta N en la línea finita se satisface también. Podrían pensar que no existe tal solución para la línea infinita, porque todas nuestras soluciones se parecen a e^{ikx_n} que da la misma norma de la amplitud en todas partes. Pero recordarán que la energía depende solamente del valor absoluto de k , por lo que otra solución, que es igualmente válida para la misma energía, sería e^{-ikx_n} . Y lo mismo es verdad para cualquier superposición de estas dos soluciones. Restándolas podemos obtener las soluciones $\sin kx_n$, que satisface el requisito de que la amplitud sea cero para $x = 0$. Aún corresponde a la energía $(E_0 - 2A \cos kb)$. Ahora bien, con una elección conveniente para el valor de k , también podemos hacer que la amplitud sea cero en x_{N+1} . Esto requiere que $(N + 1)kb$ sea un múltiplo de π , o que

$$kb = \frac{\pi}{(N + 1)} s, \quad (15.29)$$

donde s es un entero que va de 1 a N . (Tomamos solamente las k positivas, porque cada solución contiene $+k$ y $-k$; cambiar el signo de k nos vuelve a dar el mismo estado.) Para la molécula de butadieno, $N = 4$, hay cuatro estados con

$$kb = \pi/5, \quad 2\pi/5, \quad 3\pi/5 \quad \text{y} \quad 4\pi/5. \quad (15.30)$$

Podemos representar los niveles de energía usando un diagrama circular similar al del benceno. Esta vez usaremos un semicírculo dividido en cinco partes iguales como muestra la figura 15-11. El punto inferior corresponde a $s = 0$, que no da ningún estado. Lo mismo ocurre con el punto superior, que corresponde a $s = N + 1$. Los cuatro puntos restantes nos dan cuatro energías permitidas. Hay cuatro estados estacionarios, que es justamente lo que esperamos por haber comenzado con cuatro estados base. En el diagrama circular los intervalos angulares son $\pi/5$ o 36 grados. La energía más baja es de $(E_0 - 1.618A)$. (¡Ah, qué maravillas contienen las matemáticas! La regla de oro de los griegos† nos da el estado de energía más bajo de la molécula de butadieno según esta teoría.)

Ahora podemos calcular la energía de la molécula de butadieno cuando colocamos cuatro electrones. Con cuatro electrones completamos los dos niveles inferiores, cada uno con dos electrones de espín opuestos. La energía total es

$$E = 2(E_0 - 1.618A) + 2(E_0 - 0.618A) = 4(E_0 - A) - 0.472A. \quad (15.31)$$

Este resultado parece razonable. La energía es un poco menor que para dos enlaces dobles sencillos, pero la unión no es tan fuerte como en el benceno.

† La razón de los lados de un rectángulo que se puede dividir en un cuadrado y en un rectángulo semejante.

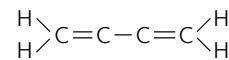


Figura 15-9. La representación del enlace de valencia de la molécula de 1,3-butadieno.

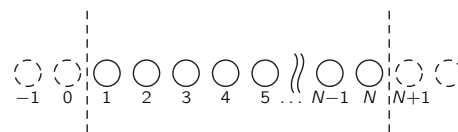


Figura 15-10. Una línea de N moléculas.

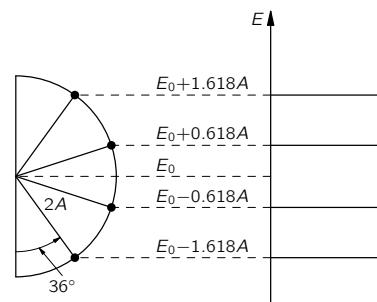


Figura 15-11. Los niveles de energía del butadieno.

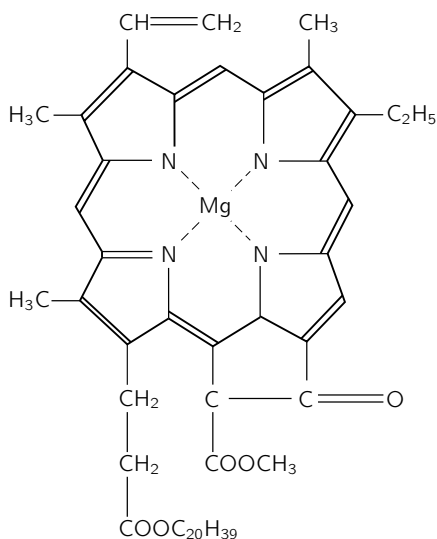


Figura 15-12. Una molécula de clorofila.

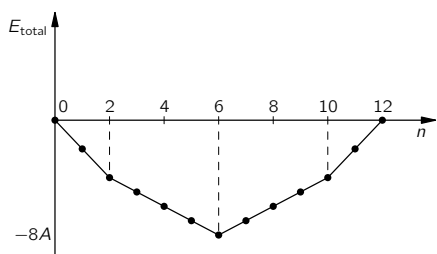


Figura 15-13. La suma de todas las energías de los electrones cuando los estados inferiores que se muestran en la figura 15-8 están ocupados por n electrones si tomamos $E_0 = 0$.

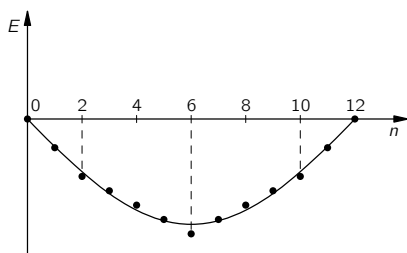


Figura 15-14. Los puntos de la figura 15-13 con una curva suave. Las moléculas con $n = 2, 6, 10$ son más estables que las demás.

De todas maneras, éste es el modo en que el químico analiza ciertas moléculas orgánicas.

El químico no sólo puede usar las energías, sino también las amplitudes de probabilidad. Al conocer las amplitudes para cada estado, y los estados que están ocupados, puede decir cuál es la probabilidad para encontrar a un electrón en cualquier lugar de la molécula. Los lugares donde es más probable que el electrón se encuentre son aptos para ser sitios reactivos en sustituciones químicas, las cuales requieren que un electrón sea compartido con algún otro grupo de átomos. Es más probable que las otras posiciones sean más reactivas en las sustituciones que tienden a ceder un electrón adicional al sistema.

Las mismas ideas que hemos estado usando pueden darnos cierta comprensión de una molécula aún tan complicada como la de clorofila, una versión de la cual se encuentra en la figura 15-12. Noten que los enlaces dobles y simples que hemos dibujado con líneas gruesas forman un largo anillo cerrado con 20 intervalos. Los electrones adicionales de los enlaces dobles pueden viajar alrededor de este anillo. Usando el método de partículas independientes, podemos obtener un conjunto completo de niveles de energía. Hay líneas de absorción fuertes debidas a transiciones entre los niveles que se encuentran en la parte visible del espectro, y que dan a esta molécula su intenso color. Moléculas complicadas similares como las xantófilas, que hacen que las hojas se vuelvan rojas, se pueden estudiar del mismo modo.

Hay una idea más que surge de la aplicación de esta teoría en química orgánica. Es probablemente la de más éxito o, por lo menos en un cierto sentido, la más exacta. Tiene que ver con la pregunta: ¿en qué situaciones se tiene un enlace químico particularmente fuerte? La respuesta es muy interesante. Tomemos, en primer lugar, el ejemplo del benceno, e imaginemos la secuencia de sucesos que tienen lugar cuando comenzamos con la molécula seis veces ionizada y añadimos más y más electrones. Estaríamos entonces pensando en varios iones de benceno —negativos o positivos—. Supongan que graficamos la energía del ion (o molécula neutra) en función del número de electrones. Si tomamos $E_0 = 0$ (ya que no sabemos cuánto vale), obtenemos la curva de la figura 15-13. Para los dos primeros electrones, la pendiente de la función es una línea recta. Para cada grupo sucesivo la pendiente aumenta, y hay una discontinuidad en la pendiente entre los grupos de electrones. La pendiente cambia justamente cuando se ha terminado de llenar un conjunto de niveles de igual energía, y se debe trasladar hacia arriba al conjunto de niveles próximo para el electrón siguiente.

La energía real del ion de benceno es en verdad bastante diferente a la de la curva de la figura 15-13, debido a las interacciones de los electrones, y a las energías electrostáticas que hemos estado ignorando. Sin embargo, estas correcciones variarán con n de un modo bastante suave. Aunque hiciéramos todas estas correcciones, la curva de energía resultante aún tendría esos saltos justo para aquellos valores de n que completan un nivel de energía determinado.

Ahora consideren una curva muy suave que se ajuste a los puntos en promedio, como la dibujada en la figura 15-14. Podemos decir que los puntos *por encima* de esta curva tienen energías “más altas de lo normal”, y que los puntos *por debajo* de ella tienen energías “más bajas de lo normal”. En general, esperaríamos que aquellas configuraciones con energías más bajas de lo normal, tuviesen una estabilidad por encima del promedio —químicamente hablando—. Hay que tener presente que las configuraciones muy por debajo de la curva siempre tienen lugar en el extremo de uno de los segmentos rectos, es decir, cuando hay electrones suficientes para completar una “capa de energía”, como se le llama—. Ésta es una predicción muy precisa de la teoría. Las moléculas —o iones— son particularmente estables (en comparación con otras configuraciones semejantes) cuando los electrones disponibles llenan justamente una capa de energía.

Esta teoría ha explicado y predicho algunos hechos químicos muy peculiares. Para tomar un ejemplo muy simple, consideren un anillo de tres. Es casi increíble que un químico pueda hacer un anillo de tres, y que sea estable, pero se ha hecho. La figura 15-15 muestra el círculo de energía para tres electrones. Ahora bien, si colocan dos electrones en el estado inferior, tienen solamente dos de los tres

electrones que necesitan. El tercer electrón se debe colocar en un nivel mucho más alto. Según nuestra discusión, esta molécula no debería ser particularmente estable, mientras que la estructura con dos electrones debería serlo. Resulta, en realidad, que la molécula neutra de trifenil ciclopropenilo es muy difícil de formar, pero el ion positivo que muestra la figura 15-16 es relativamente fácil de formar. El anillo de tres nunca es realmente fácil, porque siempre hay una tensión muy grande cuando los enlaces en una molécula orgánica forman un triángulo equilátero. Para hacer cualquier compuesto estable, se debe estabilizar la estructura de algún modo. Pero si añaden tres anillos de benceno en los vértices, se puede formar el ion positivo. (La razón por la cual es necesario añadir anillos de benceno no se conoce realmente.)

De un modo semejante, se puede analizar el anillo de cinco lados. Si dibujan el diagrama de energía, pueden ver, en una forma cualitativa, que la estructura con seis electrones debería ser una estructura especialmente estable, de modo que esa molécula debería ser más estable como ion negativo. El anillo de cinco es bien conocido y fácil de hacer, y siempre actúa como ion negativo. Igualmente, pueden verificar fácilmente que un anillo de 4 u 8 no es muy interesante, pero que un anillo de 14 o de 10 —igual que uno de 6— debería ser especialmente estable como un objeto neutro.

15-6 Otros usos para la aproximación

Hay otras dos situaciones semejantes que describiremos sólo brevemente. Al considerar la estructura de un átomo, podemos pensar que los electrones llenan capas sucesivas. La teoría de Schrödinger del movimiento del electrón se puede tratar fácilmente para un *único* electrón que se mueve en un campo “central” —uno que varía solamente con la distancia a un punto—. ¿Cómo podemos entender entonces lo que ocurre en un átomo que tiene 22 electrones?! Un modo es usando una especie de aproximación de partículas independientes. Primero calculan lo que sucede con un electrón. Así obtienen un número de niveles de energía. Colocan un electrón en el estado de energía más baja. En un modelo aproximado, pueden continuar ignorando las interacciones de los electrones y seguir llenando capas sucesivas, pero hay un modo para obtener mejores respuestas tomando en cuenta —por lo menos de una manera aproximada— el efecto de la carga eléctrica transportada por el electrón. Cada vez que se añade un electrón, calculan su amplitud para estar en lugares diferentes, y se usa esta amplitud para estimar una especie de distribución de carga esféricamente simétrica. Usan el campo de esta distribución —junto con el campo del núcleo positivo y de todos los electrones previos— para calcular los estados disponibles para el siguiente electrón. Así pueden obtener estimaciones razonablemente correctas para las energías del átomo neutro, y para diversos estados ionizados. Encontrarán que hay capas de energía, como vimos para los electrones de una molécula cíclica. Con una capa parcialmente llena, el átomo mostrará preferencia por tomar uno o más electrones adicionales, o por perder algunos y así pasar al estado más estable con una capa llena.

Esta teoría explica la maquinaria que hay detrás de las propiedades químicas fundamentales que aparecen en la tabla periódica de los elementos. Los gases nobles son aquellos elementos en los que se acaba de completar una capa, y es especialmente difícil hacerlos reaccionar. (Algunos reaccionan por supuesto —con flúor y oxígeno, por ejemplo; pero tales compuestos están muy débilmente enlazados; los llamados gases inertes son casi inertes—.) Un átomo que tiene un electrón más o un electrón menos que un gas inerte, perderá o ganará fácilmente un electrón para pasar a una condición especialmente estable (energía baja) que proviene de tener una capa completamente llena —son los elementos químicos muy activos con valencia +1 o -1—.

El otro caso se encuentra en la física nuclear. En los núcleos atómicos, los protones y los neutrones interactúan entre sí muy fuertemente. Aun así, el modelo de partículas independientes también se puede utilizar para analizar la

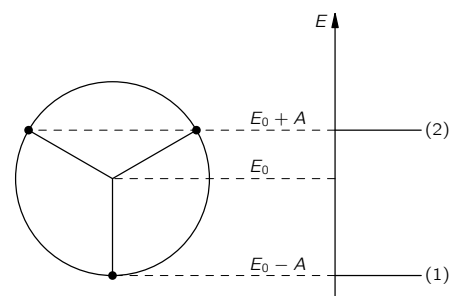


Figura 15-15. Diagrama de energía para un anillo de tres.

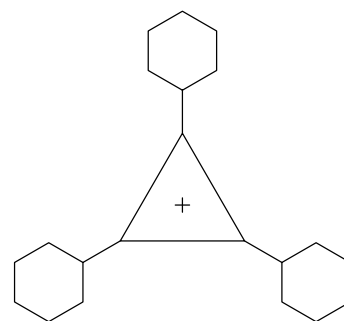


Figura 15-16. Catión trifenil ciclopropenilo.

estructura nuclear. Primero se descubrió experimentalmente que los núcleos eran especialmente estables si contenían ciertos números particulares de neutrones —específicamente 2, 8, 20, 28, 50, 82—. Los núcleos que contienen estos números de protones también son especialmente estables. Como inicialmente no hubo explicación alguna para estos números, se les llamó los “números mágicos” de la física nuclear. Es bien sabido que los protones y los neutrones interactúan entre sí fuertemente; por lo tanto, la gente se sorprendió mucho y cuando se descubrió que un modelo de partículas independientes predecía una estructura de capas, la cual producía unos pocos primeros números mágicos. El modelo suponía que cada nucleón (protón o neutrón) se movía en un potencial central, que era creado por los efectos promedio de todos los otros nucleones. Sin embargo, este modelo no pudo dar los valores correctos de los números mágicos más altos. Entonces Maria Mayer descubrió, e independientemente Jensen y sus colaboradores, que tomando el modelo de partículas independientes y añadiendo sólo una corrección por lo que se llama “interacción espín-órbita”, se podía hacer un modelo mejorado que predecía todos los números mágicos. (La interacción espín-órbita hace que la energía de un nucleón sea menor si su espín tiene la misma dirección que su momentum angular orbital de movimiento en el núcleo.) La teoría nos da aún más —su imagen de la llamada “estructura de capas” del núcleo nos ayuda a predecir ciertas características de los núcleos y de las reacciones nucleares—.

La aproximación de partículas independientes ha sido muy útil en una amplia gama de materias —desde la física del estado sólido, hasta la química, la biología y la física nuclear—. Con frecuencia es solamente una aproximación burda, pero es capaz de dar una comprensión de por qué hay condiciones especialmente estables —en capas—. Puesto que omite todas las complejidades de la interacción entre partículas individuales, no deberíamos sorprendernos de que con frecuencia falle por completo y no dé muchos detalles importantes correctamente.

La dependencia de las amplitudes con la posición

16-1 Amplitudes sobre una línea

Ahora vamos a estudiar cómo las amplitudes de probabilidad de la mecánica cuántica varían en el espacio. En algunos capítulos anteriores posiblemente hayan tenido la sensación incómoda de que hacíamos algunas omisiones. Por ejemplo, cuando hablamos de la molécula de amoníaco, elegimos describirla en términos de dos estados base. Para uno de los estados base, tomamos la situación en que el átomo de nitrógeno estaba “por encima” del plano de los tres átomos de hidrógeno, y para el otro tomamos la situación en la que el átomo de nitrógeno estaba “por debajo” del plano de los tres átomos de hidrógeno. ¿Por qué elegimos justamente estos dos estados? ¿Acaso no es posible que el átomo de nitrógeno esté a 2 ángstroms por encima del plano de los 3 átomos de hidrógeno, o a 3 o 4 ángstroms? Claro que hay muchas posiciones que el átomo de nitrógeno podría ocupar. De nuevo, cuando hablamos del ion de hidrógeno molecular, en el cual hay un electrón compartido por dos protones, imaginamos dos estados base: uno para el electrón en las cercanías del protón número uno, y el otro para el electrón en las cercanías del protón número dos. Es evidente que estábamos omitiendo muchos detalles. El electrón no está exactamente en el protón número dos, sino en sus cercanías. Podría estar en algún lugar por encima del protón, o por debajo del mismo, podría estar a la izquierda del protón, o en algún lugar a su derecha.

Evitamos intencionalmente discutir estos detalles. Dijimos que sólo estábamos interesados en ciertos aspectos del problema, por lo que imaginamos que cuando el electrón estaba en la vecindad del protón número uno, asumiría cierto estado definido con precisión. En este estado, la probabilidad para encontrar al electrón tendría una distribución bastante bien definida alrededor del protón; sin embargo, no estábamos interesados en los detalles.

Podríamos decirlo de otra manera. En nuestra discusión sobre el ion de hidrógeno molecular, elegimos una descripción aproximada al describir la situación en términos de dos estados base. En realidad, hay montones y montones de estos estados. Un electrón puede tener una condición alrededor del protón en su estado más bajo, o fundamental, pero también hay muchos estados excitados. Para cada estado excitado, la distribución del electrón alrededor del protón es diferente. Nosotros ignoramos estos estados excitados diciendo que sólo estábamos interesados en las condiciones de baja energía. Pero precisamente son estos otros estados excitados los que dan la posibilidad de diversas distribuciones del electrón alrededor del protón. Si queremos describir detalladamente al ion de hidrógeno molecular, también tenemos que tener en cuenta estos otros posibles estados base. Podríamos hacerlo de varias maneras, y una de ellas es considerando con más detalle estados en los que se describe cuidadosamente la ubicación del electrón en el espacio.

De este modo, estamos listos para considerar un procedimiento más complejo, que nos permitirá hablar en detalle sobre la posición del electrón, dando una amplitud de probabilidad para encontrar al electrón en algún lugar y en todos los lugares, en una situación dada. Esta teoría más completa sirve para mejorar las aproximaciones que hemos estado haciendo en nuestras discusiones anteriores. En cierto sentido, nuestras ecuaciones anteriores se pueden obtener como una especie de aproximación a la teoría más completa.

Ustedes se estarán preguntando por qué no comenzamos por la teoría más completa, y hacíamos las aproximaciones a medida que avanzábamos. Creímos que

16-1 Amplitudes sobre una línea

16-2 La función de onda

16-3 Estados de momentum definido

16-4 Normalización de los estados en x

16-5 La ecuación de Schrödinger

16-6 Niveles de energía cuantizados

sería mucho más fácil que ustedes alcanzaran una comprensión de la maquinaria básica de la mecánica cuántica, comenzando con las aproximaciones de dos estados y avanzando gradualmente hasta la teoría más completa, en vez de enfocar el tema en sentido opuesto. Por esta razón, nuestro enfoque del tema aparece en orden inverso al que encontrarán en muchos libros.

A medida que nos adentremos en el tema de este capítulo, observarán que estamos rompiendo una regla que siempre seguíamos antes. Siempre que tratábamos cualquier tema, intentábamos dar una descripción más o menos completa de la física del problema —mostrándoles tanto como podíamos hacia dónde nos llevaban los conceptos—. Tratábamos de describir las consecuencias generales de una teoría, así como algunos detalles específicos, de modo que pudieran ver a dónde llevaría la teoría. Ahora vamos a romper esta regla; describiremos cómo se puede hablar de las amplitudes de probabilidad en el espacio, y les mostraremos las ecuaciones diferenciales que éstas satisfacen. No tendremos tiempo para continuar y discutir muchas de las implicaciones evidentes que provienen de la teoría. De hecho, ni siquiera podremos llegar lo suficientemente lejos como para relacionar esta teoría con algunas de las formulaciones aproximadas que usamos anteriormente —por ejemplo, con la molécula de hidrógeno o con la de amoníaco—. En esta ocasión, debemos dejar nuestro trabajo abierto y sin terminar. Nos estamos aproximando al fin de nuestro curso y debemos darnos por satisfechos al tratar de darles una introducción a los conceptos generales, indicándoles las conexiones entre lo que hemos estado describiendo y alguna de las otras maneras de enfocar el tema de la mecánica cuántica. Esperamos poder darles los conceptos suficientes como para que puedan seguir por su cuenta, y al leer libros puedan aprender sobre muchas de las implicaciones de las ecuaciones que vamos a describir. Después de todo, tenemos que dejar algo para el futuro.

Revisemos una vez más lo que hemos encontrado sobre cómo un electrón se puede mover a lo largo de una línea de átomos. Cuando un electrón tiene una amplitud para saltar de un átomo al siguiente, hay estados de energía definida, en los que la amplitud de probabilidad para encontrar el electrón está distribuida a lo largo de la red en forma de onda viajera. Para longitudes de ondas grandes —para valores pequeños del número de onda k — la energía del estado es proporcional al cuadrado del número de onda. En el caso de una red cristalina con espaciamiento b , en la cual la amplitud por unidad de tiempo para que el electrón salte de un átomo al siguiente es $iA/\hbar$, la energía del estado está relacionada con k (para kb pequeño) mediante

$$E = Ak^2b^2 \quad (16.1)$$

(véase sección 13-2). También vimos que grupos de esas ondas con energías parecidas formarían un paquete de ondas, el cual se comportaría como una partícula clásica con una masa efectiva m_{ef} dada por:

$$m_{\text{ef}} = \frac{\hbar^2}{2Ab^2}. \quad (16.2)$$

Como las ondas de amplitud de probabilidad en un cristal se comportan como una partícula, sería de esperar que la descripción cuántica general de una partícula muestre el mismo tipo de comportamiento ondulatorio que observamos en la red cristalina. Supongan que consideráramos una red lineal, e imaginen que el espaciamiento b de la red se hiciera más y más pequeño. En el límite, estaríamos considerando el caso en que el electrón podría estar en cualquier parte a lo largo de la línea. Habríamos pasado a una distribución continua de amplitudes de probabilidad. Tendríamos la amplitud para encontrar a un electrón en cualquier parte, a lo largo de la línea. Ésta sería una manera para describir el movimiento de un electrón en el vacío. En otras palabras, si imaginamos que se puede etiquetar al espacio con una infinidad de puntos muy juntos, y podemos obtener las ecuaciones que relacionan las amplitudes en un punto con las amplitudes en puntos vecinos, tendremos las leyes cuánticas para el movimiento de un electrón en el espacio.

Comencemos recordando algunos principios generales de la mecánica cuántica. Supongan que tenemos una partícula, la cual puede existir en diversas condiciones

en un sistema cuántico. A cualquier condición en la que se pueda encontrar a un electrón, la llamamos “estado”, y la etiquetamos con un vector de estado, digamos $|\phi\rangle$. Otra condición se designaría con otro vector de estado, por ejemplo $|\psi\rangle$. Introducimos luego el concepto de estados base. Decimos que hay un conjunto de estados $|1\rangle, |2\rangle, |3\rangle, |4\rangle$, etc., que tienen las siguientes propiedades. Primero, todos estos estados son completamente diferentes —decimos que son ortogonales—. Entendemos por esto que, para cualquier par de estados base $|i\rangle$ y $|j\rangle$, la amplitud $\langle i | j \rangle$ para que un electrón, que se sabe está en el estado $|i\rangle$, esté también en el estado $|j\rangle$ es igual a cero —a no ser, por supuesto, que $|i\rangle$ y $|j\rangle$ designen al mismo estado—. Representamos esto de manera simbólica mediante

$$\langle i | j \rangle = \delta_{ij}. \quad (16.3)$$

Recordarán que $\delta_{ij} = 0$ si i y j son diferentes, y que $\delta_{ij} = 1$ si i y j son iguales.

Segundo, los estados base $|i\rangle$ deben ser un conjunto completo, de modo que cualquier estado se pueda describir en términos de ellos. Esto es, cualquier estado $|\phi\rangle$ se puede describir completamente dando todas las amplitudes $\langle i | \phi \rangle$ para que una partícula en el estado $|\phi\rangle$, se encuentre también en el estado $|i\rangle$. De hecho, el vector de estado $|\phi\rangle$ es igual a la suma de los estados base multiplicados cada uno por un coeficiente, que es la amplitud para que el estado $|\phi\rangle$ esté también en el estado $|i\rangle$:

$$|\phi\rangle = \sum_i |i\rangle \langle i | \phi \rangle. \quad (16.4)$$

Finalmente, si consideramos dos estados $|\phi\rangle$ y $|\psi\rangle$ cualesquiera, se puede encontrar la amplitud para que el estado $|\psi\rangle$ esté también en el estado $|\phi\rangle$, proyectando primero al estado $|\psi\rangle$ sobre los estados base, y luego proyectando desde cada estado base al estado $|\phi\rangle$. Escribimos esto de la siguiente manera:

$$\langle \phi | \psi \rangle = \sum_i \langle \phi | i \rangle \langle i | \psi \rangle. \quad (16.5)$$

La suma se extiende, naturalmente, a todo el conjunto de estados base $|i\rangle$.

En el capítulo 13, cuando tratábamos lo que ocurre con un electrón colocado en un arreglo lineal de átomos, elegimos un conjunto de estados base, en el que el electrón estaba localizado en uno u otro de los átomos de la línea. El estado de base $|n\rangle$ representaba la situación en que el electrón estaba localizado en el átomo número “ n ”. (Por supuesto, no tiene importancia el que llamáramos $|n\rangle$ a nuestros estados base en vez de $|i\rangle$.) Un poco después, encontramos que era conveniente designar a los estados base por medio de la coordenada x_n del átomo, en vez de hacerlo con el número del átomo en el arreglo. El estado $|x_n\rangle$ es simplemente otra manera de escribir al estado $|n\rangle$. Por lo tanto, siguiendo las reglas generales, un estado cualquiera, $|\psi\rangle$ por ejemplo, queda descrito dando las amplitudes para que un electrón en el estado $|\psi\rangle$ esté también en uno de los estados $|x_n\rangle$. Por conveniencia, hemos escogido al símbolo C_n para designar estas amplitudes,

$$C_n = \langle x_n | \psi \rangle. \quad (16.6)$$

Como los estados base están asociados con una ubicación a lo largo de la línea, podemos considerar a la amplitud C_n como función de la coordenada x y escribirla en la forma $C(x_n)$. En general, las amplitudes $C(x_n)$ variarán con el tiempo, y en consecuencia serán también funciones de t . Por lo general, no nos molestaremos en mostrar explícitamente esta dependencia.

En el capítulo 13 propusimos que las amplitudes $C(x_n)$ deberían variar en el tiempo de una manera descrita por la ecuación hamiltoniana [ecuación (13.3)]. En nuestra nueva notación, esta ecuación es

$$i\hbar \frac{\partial C(x_n)}{\partial t} = E_0 C(x_n) - AC(x_n + b) - AC(x_n - b). \quad (16.7)$$

Los dos últimos términos del lado derecho representan el proceso en el que un electrón que está en el átomo $(n + 1)$ o en el $(n - 1)$ pueda llegar al átomo n .

Encontramos que la ecuación (16.7) tiene soluciones que corresponden a estados de energía definida, que escribimos en la forma

$$C(x_n) = e^{-iEt/\hbar} e^{ikx_n}. \quad (16.8)$$

Para los estados de baja energía, las longitudes de onda son grandes (k es pequeño), y la energía está relacionada con k mediante

$$E = (E_0 - 2A) + Ak^2b^2, \quad (16.9)$$

o, escogiendo nuestro cero de energía de modo que $(E_0 - 2A) = 0$, la energía está dada por la ecuación (16.1).

Veamos lo que ocurriría si hiciéramos que el espaciamiento de la red b tendiera a cero, manteniendo fijo el número de onda k . Si sólo se hiciera eso, el último término de la ecuación (16.9) tendería a cero, y nos quedaríamos sin física. Pero supongan que A y b varían conjuntamente, de modo que cuando b tiende a cero, el producto Ab^2 se mantiene constante[†] —usando la ecuación (16.2), escribiremos Ab^2 como la constante $\hbar^2/2m_{\text{ef}}$ —. En estas circunstancias, la ecuación (16.9) no cambiaría, pero ¿qué le ocurriría a la ecuación diferencial (16.7)?

Primero, volveremos a escribir la ecuación (16.7) en la forma

$$i\hbar \frac{\partial C(x_n)}{\partial t} = (E_0 - 2A)C(x_n) + A[2C(x_n) - C(x_n + b) - C(x_n - b)]. \quad (16.10)$$

Con nuestra elección de E_0 , el primer término desaparece. A continuación, podemos pensar en una función continua $C(x)$ que pasa suavemente por los valores apropiados $C(x_n)$ en cada x_n . Cuando el espaciamiento b tiende a cero, los puntos x_n se acercan más y más, y [si mantenemos razonablemente suave la variación de $C(x)$] la cantidad entre paréntesis es proporcional a la segunda derivada de $C(x)$. Podemos escribir —como pueden ver haciendo una expansión de Taylor de cada término— la igualdad

$$2C(x) - C(x + b) - C(x - b) \approx -b^2 \frac{\partial^2 C(x)}{\partial x^2}. \quad (16.11)$$

Entonces, en el límite cuando b tiende a cero, y manteniendo b^2A igual a K , la ecuación (16.7) tiende a

$$i\hbar \frac{\partial C(x)}{\partial t} = -\frac{\hbar^2}{2m_{\text{ef}}} \frac{\partial^2 C(x)}{\partial x^2}. \quad (16.12)$$

Tenemos una ecuación que dice que la rapidez de variación temporal de $C(x)$ —la amplitud para encontrar al electrón en x — depende de la amplitud para encontrar al electrón en puntos vecinos de una manera proporcional a la segunda derivada de la amplitud respecto a la posición.

La ecuación cuántica correcta para el movimiento de un electrón en el espacio libre fue descubierta primero por Schrödinger. Para el movimiento en una línea recta, tiene exactamente la forma de la ecuación (16.12) si reemplazamos m_{ef} por m , masa del electrón en el espacio libre. La ecuación de Schrödinger para el movimiento a lo largo de una línea en el espacio libre es

$$i\hbar \frac{\partial C(x)}{\partial t} = -\frac{\hbar^2}{2m} \frac{\partial^2 C(x)}{\partial x^2}. \quad (16.13)$$

No tenemos la intención de hacerles creer que hemos deducido la ecuación de Schrödinger, sino que sólo deseamos mostrarles una manera de pensar en ella. Cuando Schrödinger la escribió por primera vez, dio una especie de deducción basada en algunos argumentos heurísticos y en algunas brillantes conjeturas intuitivas. Algunos de los argumentos que usó hasta eran falsos, pero no importa; lo único importante es que la ecuación final da una descripción correcta de

[†] Pueden imaginar que, a medida que los puntos x_n se acercan, aumentará la amplitud A para saltar de $x_{n\pm 1}$ a x_n .

la naturaleza. El propósito de nuestra discusión es, entonces, simplemente mostrarles que la ecuación cuántica fundamental correcta (16.13) tiene la misma forma que se obtiene como caso límite de un electrón moviéndose a lo largo de una línea de átomos. Lo que indica que podemos considerar que la ecuación diferencial (16.13) describe la difusión de una amplitud de probabilidad desde un punto hasta el siguiente a lo largo de la línea. Es decir, si un electrón tiene cierta amplitud para estar en un punto, un poco más tarde tendrá una amplitud para estar en los puntos vecinos. En realidad, la ecuación se parece algo a las ecuaciones de difusión que usamos en el volumen I. Pero hay una diferencia importante: el coeficiente imaginario delante de la derivada temporal hace que el comportamiento sea completamente diferente al de la difusión ordinaria, como la que tendrían para un gas esparciéndose a lo largo de un tubo delgado. La difusión ordinaria da lugar a soluciones exponenciales reales, mientras que las soluciones de la ecuación (16.13) son ondas complejas.

16-2 La función de onda

Ahora que tienen una idea sobre cómo van a ser las cosas, queremos volver al principio y estudiar el problema de describir el movimiento de un electrón a lo largo de una línea, sin tener que considerar estados relacionados con átomos en una red. Queremos volver al principio y ver qué conceptos tenemos que usar si queremos describir el movimiento de una partícula libre en el espacio. Como estamos interesados en el comportamiento de una partícula en un medio continuo, manejaremos un número infinito de estados posibles, y, como verán, los conceptos que hemos desarrollado para tratar un número finito de estados necesitarán algunas modificaciones técnicas.

Comencemos diciendo que el vector de estado $|x\rangle$ representa un estado en el que una partícula está ubicada precisamente en la coordenada x . Para cada valor de x a lo largo de la línea —por ejemplo 1.73, 9.67 o 10.00— hay un estado correspondiente. Tomaremos estos estados $|x\rangle$ como nuestros estados base y, si incluimos todos los puntos de la línea, tendremos un conjunto completo para el movimiento en una dimensión. Supongamos ahora que tenemos un estado diferente, por ejemplo $|\psi\rangle$, en el que un electrón está distribuido de cierto modo a lo largo de la línea. Una manera para describir este estado es dar todas las amplitudes para que el electrón se encuentre también en cada uno de los estados base $|x\rangle$. Debemos dar un conjunto infinito de amplitudes, una para cada valor de x . Escribiremos estas amplitudes en la forma $\langle x|\psi\rangle$. Cada una de estas amplitudes es un número complejo y, como hay uno de estos números complejos para cada valor de x , la amplitud $\langle x|\psi\rangle$ es en realidad una función de x . También las escribiremos como $C(x)$,

$$C(x) \equiv \langle x|\psi\rangle. \quad (16.14)$$

Ya hemos considerado estas amplitudes que varían en forma continua con las coordenadas, cuando hablamos de las variaciones temporales de la amplitud en el capítulo 7. Demostramos ahí, por ejemplo, que una partícula de momentum definido debería tener una variación particular de su amplitud en el espacio. Si una partícula tiene un momentum definido p , y una energía definida E correspondiente, la amplitud para encontrarse en cualquier posición x tendría el aspecto siguiente:

$$\langle x|\psi\rangle = C(x) \propto e^{+ipx/\hbar}. \quad (16.15)$$

Esta ecuación expresa un principio general muy importante de la mecánica cuántica, el cual conecta los estados base correspondientes a diferentes posiciones en el espacio, con otro sistema de estados base —todos los estados de momentum definido—. Los estados de momentum definido son a menudo más convenientes para ciertos tipos de problemas que los estados en x . Por supuesto, cualquiera de los dos conjuntos de estados base es igualmente aceptable para describir una situación cuántica. Más adelante hablaremos sobre la conexión entre ellos. Por el

momento, nos concentraremos en nuestro estudio de una descripción en términos de los estados $|x\rangle$.

Antes de seguir, queremos hacer un pequeño cambio de notación, que esperamos no cause demasiada confusión. La función $C(x)$, definida en la ecuación (16.14), tendrá naturalmente una forma que depende del estado particular $|\psi\rangle$ que se considere. Tenemos que indicar esto de alguna manera. Por ejemplo, podríamos especificar sobre qué función $C(x)$ estamos hablando por medio de un subíndice, digamos $C_\psi(x)$. Aunque ésta sería una notación totalmente satisfactoria, es un poco incómoda y no es la que encontrarán en la mayoría de los libros. Casi todo el mundo omite simplemente la letra C y usa el símbolo ψ para definir a la función

$$\psi(x) \equiv C_\psi(x) = \langle x | \psi \rangle. \quad (16.16)$$

Como ésta es la notación que usa todo el mundo, es mejor que se acostumbren a ella para que no se asusten cuando la encuentren en alguna otra parte. No obstante, recuerden que ahora estaremos usando a ψ en dos maneras diferentes. En la ecuación (16.14), ψ es una etiqueta que le hemos puesto a un estado físico particular del electrón. Por otra parte, en el lado izquierdo de la ecuación (16.16), el símbolo ψ está usado para definir una función matemática de x , que es igual a la amplitud que hay que asociar con cada punto x a lo largo de la línea. Esperamos que no haya confusión una vez que se acostumbren a la idea. A propósito, la función $\psi(x)$ por lo común se llama “la función de onda” —porque la mayoría de las veces tiene la forma de una onda compleja en sus variables—.

Ya que hemos definido a $\psi(x)$ como la amplitud para que un electrón en el estado ψ se encuentre en la posición x , nos gustaría interpretar al cuadrado de la norma de ψ como la probabilidad para encontrar a un electrón en la posición x . Desafortunadamente, la probabilidad para encontrar una partícula exactamente en una posición determinada es cero. Generalmente, el electrón estará diseminado en cierta región de la línea y, como en cualquier trozo de la línea, hay un número infinito de puntos, la probabilidad para que esté en cualquiera de ellos no puede ser un número finito. Únicamente podemos describir la probabilidad para encontrar a un electrón por medio de una *distribución de probabilidad*[†] que da la probabilidad *relativa* para encontrar al electrón en diversas ubicaciones aproximadas a lo largo de la línea. Designemos con $\text{prob}(x, \Delta x)$ a la probabilidad para encontrar al electrón en un pequeño intervalo Δx ubicado cerca de x . Si vamos a una escala suficientemente pequeña en cualquier situación física, la probabilidad variará suavemente de punto a punto, y la probabilidad para encontrar al electrón en cualquier segmento pequeño finito Δx será proporcional a Δx . Podemos modificar nuestras definiciones para tener en cuenta esto.

Podemos considerar que la amplitud $\langle x | \psi \rangle$ representa una especie de “densidad de amplitud” para todos los estados base $|x\rangle$ que hay en una pequeña región. Como la probabilidad para encontrar a un electrón en un pequeño intervalo Δx en x debe ser proporcional al intervalo Δx , escogemos nuestra definición de $\langle x | \psi \rangle$ de modo que cumpla la siguiente relación:

$$\text{prob}(x, \Delta x) = |\langle x | \psi \rangle|^2 \Delta x.$$

La amplitud $\langle x | \psi \rangle$ es, por lo tanto, proporcional a la amplitud para que un electrón en el estado ψ se encuentre en el estado base x , y la constante de proporcionalidad se elige de modo que el cuadrado de la norma de la amplitud $\langle x | \psi \rangle$ dé la *densidad de probabilidad* para encontrar a un electrón en cualquier región pequeña. En forma equivalente, podemos escribir,

$$\text{prob}(x, \Delta x) = |\psi(x)|^2 \Delta x. \quad (16.17)$$

Ahora tendremos que modificar algunas de las ecuaciones que obtuvimos antes, para hacerlas compatibles con esta nueva definición de amplitud de probabilidad. Supongan que tenemos a un electrón en el estado $|\psi\rangle$ y queremos conocer la

[†] Para una discusión sobre las distribuciones de probabilidad, véase la sección 6-4 del volumen I.

amplitud para encontrarlo en un estado diferente $|\phi\rangle$ que puede corresponder a una distribución diferente del electrón. Cuando hablamos de un conjunto finito de estados discretos, habríamos usado la ecuación (16.5). Antes de modificar nuestra definición de las amplitudes habríamos escrito

$$\langle\phi|\psi\rangle = \sum_{\text{toda } x} \langle\phi|x\rangle\langle x|\psi\rangle. \quad (16.18)$$

Ahora bien, si estas dos amplitudes están normalizadas como indicamos antes, una suma de todos los estados dentro de una pequeña región de x sería equivalente a multiplicar por Δx , y la suma sobre todos los valores de x se convierte simplemente en una integral. Con nuestras definiciones modificadas, la forma correcta es

$$\langle\phi|\psi\rangle = \int_{\text{toda } x} \langle\phi|x\rangle\langle x|\psi\rangle dx. \quad (16.19)$$

La amplitud $\langle x|\psi\rangle$ es lo que ahora estamos llamando $\psi(x)$ y, análogamente, a la amplitud $\langle x|\phi\rangle$ la estamos representando por $\phi(x)$. Recordando que $\langle\phi|x\rangle$ es el complejo conjugado de $\langle x|\phi\rangle$, podemos escribir la ecuación (16.19) en la forma

$$\langle\phi|\psi\rangle = \int \phi^*(x)\psi(x) dx. \quad (16.20)$$

Con nuestras nuevas definiciones, todo se hace con las mismas fórmulas que antes si se reemplaza siempre un signo de suma por una integral sobre x .

Debemos mencionar una limitación a lo que hemos estado diciendo. Todo conjunto apropiado de estados base debe ser completo para poder usarlo en una descripción adecuada de lo que está pasando. Para un electrón en una dimensión, en realidad no es suficiente especificar únicamente los estados base $|x\rangle$, pues en cada uno de estos estados el electrón puede tener espín hacia arriba o hacia abajo. Una manera de obtener un conjunto completo es tomar dos conjuntos de estados en x , uno para espín hacia arriba y otro para espín hacia abajo. Sin embargo, por ahora no nos preocuparemos de esas complicaciones.

16-3 Estados de momentum definido

Supongan que tenemos un electrón en un estado $|\psi\rangle$ que está descrito por la amplitud de probabilidad $\langle x|\psi\rangle = \psi(x)$. Sabemos que esto representa un estado en el que el electrón está diseminado a lo largo de la línea con cierta distribución, de modo que la probabilidad para encontrar al electrón en un pequeño intervalo dx en la posición x es simplemente

$$\text{prob}(x, dx) = |\psi(x)|^2 dx.$$

¿Qué podemos decir sobre el momentum de este electrón? Podríamos preguntar ¿cuál es la probabilidad para que este electrón tenga momentum p ? Comencemos calculando la amplitud para que el estado $|\psi\rangle$ esté en otro estado $|\text{mom } p\rangle$, que definimos como estado con el momentum definido p . Podemos encontrar esta amplitud empleando nuestra ecuación básica para la resolución de amplitudes, ecuación (16.19). En función del estado $|\text{mom } p\rangle$

$$\langle\text{mom } p|\psi\rangle = \int_{x=-\infty}^{+\infty} \langle\text{mom } p|x\rangle\langle x|\psi\rangle dx. \quad (16.21)$$

Y la probabilidad para que encontremos al electrón con momentum p debe estar dada en términos del cuadrado de la norma de esta amplitud. Sin embargo, tenemos de nuevo un pequeño problema con la normalización. En general, sólo podemos preguntar sobre la probabilidad para encontrar a un electrón con momentum dentro de un pequeño intervalo dp alrededor del momentum p . La probabilidad para que el momentum sea exactamente un cierto valor p debe ser cero (a no ser que ocurra que el estado $|\psi\rangle$ sea un estado de momentum definido).

Sólo si preguntamos por la probabilidad para encontrar el momentum dentro de un pequeño intervalo dp alrededor del momentum p obtendremos una probabilidad finita. Hay varias maneras de ajustar las normalizaciones. Escogeremos una que creemos es la más conveniente, aunque en este momento no parezca así.

Tomamos nuestra normalización de modo que la probabilidad esté relacionada con la amplitud mediante

$$\text{prob}(p, dp) = |\langle \text{mom } p | \psi \rangle|^2 \frac{dp}{2\pi\hbar}. \quad (16.22)$$

Con esta definición, la normalización de la amplitud $\langle \text{mom } p | x \rangle$ queda determinada. La amplitud $\langle \text{mom } p | x \rangle$ es, por supuesto, el complejo conjugado de la amplitud $\langle x | \text{mom } p \rangle$, que es precisamente la que hemos escrito en la ecuación (16.15). Con la normalización que hemos elegido, resulta que la constante apropiada de proporcionalidad que lleva la exponencial es simplemente 1. Es decir,

$$\langle \text{mom } p | x \rangle = \langle x | \text{mom } p \rangle^* = e^{-ipx/\hbar}. \quad (16.23)$$

La ecuación (16.21) se convierte entonces en

$$\langle \text{mom } p | \psi \rangle = \int_{-\infty}^{+\infty} e^{-ipx/\hbar} \langle x | \psi \rangle dx. \quad (16.24)$$

Esta ecuación, junto con la ecuación (16.22), nos permite encontrar la distribución de momentum para cualquier estado $|\psi\rangle$.

Examinemos un ejemplo particular —por ejemplo, uno en el que un electrón está localizado en cierta región alrededor de $x = 0$ —. Supongan que tomamos una función de onda que tiene la forma siguiente:

$$\psi(x) = K e^{-x^2/4\sigma^2}. \quad (16.25)$$

La distribución de probabilidad en x para esta función de onda es el cuadrado de la norma, o sea

$$\text{prob}(x, dx) = P(x) dx = K^2 e^{-x^2/2\sigma^2} dx. \quad (16.26)$$

La función de densidad de probabilidad $P(x)$ es la curva gaussiana que muestra la figura 16-1. La mayor parte de la probabilidad está concentrada entre $x = +\sigma$ y $x = -\sigma$. Decimos que la “anchura a media altura” de la curva es σ . (Para ser más precisos, σ es igual a la media cuadrática de la coordenada x de algo diseminado con esta distribución.) Normalmente elegimos a la constante K de modo que la densidad de probabilidad $P(x)$ no sólo sea *proporcional* a la probabilidad por unidad de longitud en x para encontrar el electrón, sino que tenga una escala tal que $P(x)\Delta x$ sea *igual* a la probabilidad para encontrar al electrón en Δx cerca de x . La constante K que hace esto se puede encontrar pidiendo que $\int_{-\infty}^{+\infty} P(x) dx = 1$, ya que la probabilidad para encontrar al electrón en algún lado debe ser uno. Así obtenemos $K = (2\pi\sigma^2)^{-1/4}$. [Hemos usado que $\int_{-\infty}^{+\infty} e^{-t^2} dt = \sqrt{\pi}$; véase el volumen I, página 40-7.]

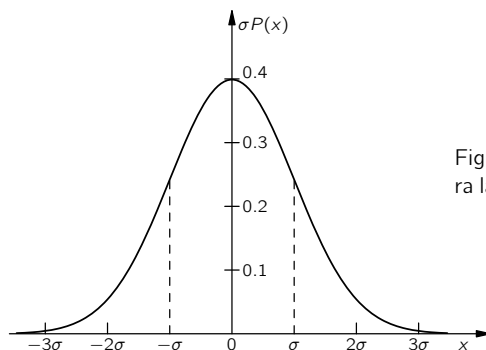


Figura 16-1. La densidad de probabilidad para la función de onda de la ecuación (16.25).

Ahora encontremos la distribución de momentum. Designemos con $\phi(p)$ a la amplitud para encontrar al electrón con el momentum p ,

$$\phi(p) \equiv \langle \text{mom } p | \psi \rangle. \quad (16.27)$$

Sustituyendo la ecuación (16.25) en la ecuación (16.24) obtenemos

$$\phi(p) = \int_{-\infty}^{+\infty} e^{-ipx/\hbar} \cdot K e^{-x^2/4\sigma^2} dx. \quad (16.28)$$

La integral también se puede escribir en la forma

$$K e^{-p^2\sigma^2/\hbar^2} \int_{-\infty}^{+\infty} e^{-(1/4\sigma^2)(x+2ip\sigma^2/\hbar)^2} dx. \quad (16.29)$$

Hacemos ahora la sustitución $u = x + 2ip\sigma^2/\hbar$, por lo que la integral es

$$\int_{-\infty}^{+\infty} e^{-u^2/4\sigma^2} du = 2\sigma\sqrt{\pi}. \quad (16.30)$$

(Probablemente los matemáticos pondrían objeciones a la manera en que llegamos aquí, no obstante, el resultado es correcto.)

$$\phi(p) = (8\pi\sigma^2)^{1/4} e^{-p^2\sigma^2/\hbar^2}. \quad (16.31)$$

Obtenemos el resultado interesante de que la función de amplitud en p tiene precisamente la misma forma matemática que la función de amplitud en x ; sólo el ancho de la gaussiana es diferente. Podemos escribirla en la forma

$$\phi(p) = (\eta^2/2\pi\hbar^2)^{-1/4} e^{-p^2/4\eta^2}, \quad (16.32)$$

donde la anchura a media altura η de la función de distribución en p está relacionada con la anchura a media altura σ de la distribución en x mediante

$$\eta = \frac{\hbar}{2\sigma}. \quad (16.33)$$

Nuestro resultado dice: si hacemos el ancho de la distribución en x muy pequeño, haciendo σ pequeña, η se hace grande, y la distribución en p se ensancha mucho. Inversamente: si tenemos una distribución estrecha en p , ésta debe corresponder a una distribución ancha en x . Si queremos, podemos considerar que η y σ son una medida de la incertidumbre en la localización del momentum y de la posición del electrón en el estado que estamos estudiando. Si las llamamos Δp y Δx , respectivamente, la ecuación (16.33) se convierte en

$$\Delta p \Delta x = \frac{\hbar}{2}. \quad (16.34)$$

Un resultado muy interesante: es posible demostrar que para cualquier otra forma de distribución en x o en p , el producto $\Delta p \Delta x$ no puede ser menor de lo que hemos encontrado aquí. La distribución gaussiana da el valor más pequeño posible para el producto de las medias cuadráticas de los anchos. En general, podemos decir

$$\Delta p \Delta x \geq \frac{\hbar}{2}. \quad (16.35)$$

Éste es un enunciado cuantitativo del principio de incertidumbre de Heisenberg, el cual ya hemos discutido cualitativamente muchas veces. Usualmente hemos hecho el enunciado aproximado de que el valor mínimo del producto $\Delta p \Delta x$ es del mismo orden que $\hbar$.

16-4 Normalización de los estados en x

Volvemos ahora a la discusión de las modificaciones en nuestras ecuaciones básicas, que son necesarias para tratar un medio continuo de estados base. Cuando tenemos un número finito de estados discretos, el conjunto de estados base debe satisfacer la siguiente condición fundamental:

$$\langle i | j \rangle = \delta_{ij}. \quad (16.36)$$

Si una partícula está en un estado base, la amplitud para estar en otro estado base es 0. Eligiendo una normalización apropiada, hemos definido a la amplitud $\langle i | i \rangle$ igual a 1. La ecuación (16.36) describe estas dos condiciones. Ahora queremos ver cómo hay que modificar esta relación cuando usamos los estados base $|x\rangle$ de una partícula sobre una línea. Si se sabe que la partícula está en uno de los estados base $|x\rangle$, ¿cuál es la amplitud para que esté en otro estado base $|x'\rangle$? Si x y x' son dos posiciones diferentes a lo largo de la línea, la amplitud $\langle x | x' \rangle$ es con certeza 0, de modo que es consistente con la ecuación (16.36). Pero si x y x' son iguales, la amplitud $\langle x | x' \rangle$ no será 1, debido al mismo problema de siempre de la normalización. Para ver cómo tenemos que remendar las cosas, volvamos a la ecuación (16.19), y apliquémosla al caso especial en que el estado $|\phi\rangle$ es justamente el estado base $|x'\rangle$. Tendríamos entonces

$$\langle x' | \psi \rangle = \int \langle x' | x \rangle \psi(x) dx. \quad (16.37)$$

Ahora la amplitud $\langle x | \psi \rangle$ es precisamente lo que hemos estado llamando función $\psi(x)$. Análogamente la amplitud $\langle x' | \psi \rangle$, como se refiere al mismo estado $|\psi\rangle$, es la misma función de la variable x' , o sea $\psi(x')$. Por consiguiente, podemos reescribir la ecuación (16.37) en la forma

$$\psi(x') = \int \langle x' | x \rangle \psi(x) dx. \quad (16.38)$$

Esta ecuación debe ser válida para cualquier estado $|\psi\rangle$ y, por lo tanto, para cualquier función arbitraria $\psi(x)$. Este requisito debería determinar completamente la naturaleza de la amplitud $\langle x | x' \rangle$ —que es, por supuesto, simplemente una función que depende de x y de x' —.

Nuestro problema es ahora encontrar una función $f(x, x')$ que multiplicada por $\psi(x)$, e integrada sobre toda x dé justamente la cantidad $\psi(x')$. ¡Y resulta que no hay ninguna función matemática que haga esto! Al menos nada de lo que por lo general entendemos por “función”.

Supongan que tomamos para x' el número especial 0 y definimos la amplitud $\langle 0 | x \rangle$ como cierta función de x , digamos $f(x)$. Así, la ecuación (16.38) sería:

$$\psi(0) = \int f(x) \psi(x) dx. \quad (16.39)$$

¿Qué tipo de función $f(x)$ posiblemente podría satisfacer esta ecuación? Como la integral no debe depender de los valores que toma $\psi(x)$ para valores de x diferentes de 0, $f(x)$ debe ser evidentemente 0 para todo valor de x excepto para 0. Pero si $f(x)$ es 0 en todas partes, también la integral será 0 y no se cumplirá con la ecuación (16.39). Así pues, tenemos una situación imposible: queremos una función que sea 0 en todas partes, excepto en un punto y que dé, no obstante, una integral distinta de cero. Como no podemos encontrar una función que haga esto, la salida más fácil es simplemente decir que la función $f(x)$ está *definida* por la ecuación (16.39). Es decir, que $f(x)$ es la función que hace que la ecuación (16.39) sea correcta. Esta función fue inventada por Dirac y lleva su nombre. La escribimos $\delta(x)$. Todo lo que decimos es que la función $\delta(x)$ tiene la extraña propiedad de que si se usa como $f(x)$ en la ecuación (16.39), la integral elige el valor que toma $\psi(x)$ cuando x es igual a 0, y como la integral debe ser

independiente de $\psi(x)$ para todos los valores de x distintos de 0, la función $\delta(x)$ debe ser 0 en todas partes, excepto en $x = 0$. Resumiendo, escribimos

$$\langle 0 | x \rangle = \delta(x), \quad (16.40)$$

donde $\delta(x)$ está definida por

$$\psi(0) = \int \delta(x) \psi(x) dx. \quad (16.41)$$

Observen lo que ocurre si en la ecuación (16.41) usamos como ψ la función especial “1”. En ese caso tenemos el resultado

$$1 = \int \delta(x) dx. \quad (16.42)$$

Es decir, la función $\delta(x)$ tiene la propiedad de que es 0 en todas partes excepto en $x = 0$, pero tiene una integral finita igual a la unidad. Tenemos que imaginar que la función $\delta(x)$ tiene un infinito tan fantástico en un punto, de manera que el área total resulta igual a uno.

Una manera de imaginar el aspecto que tiene la función δ de Dirac, es considerar una sucesión de rectángulos —o cualquier otra función con un pico que les agrade— que se hacen más y más angostos y más y más altos, siempre de área uno, como se muestra en la figura 16-2. La integral de esta función de $-\infty$ a $+\infty$ siempre es 1. Si la multiplican por cualquier función $\psi(x)$ e integran el producto, obtienen algo que es aproximadamente el valor de la función en $x = 0$, y la aproximación se hace cada vez mejor cuando usan rectángulos más angostos. Si quieren, pueden imaginar a la función δ en términos de este tipo de proceso límite. Sin embargo, lo único importante es que la función δ está definida de modo tal que la ecuación (16.41) sea válida para toda función $\psi(x)$ posible. Esto define de manera única a la función δ . Sus propiedades son entonces las que hemos descrito.

Si cambiamos el argumento de la función δ de x a $x - x'$, las relaciones correspondientes son

$$\begin{aligned} \delta(x - x') &= 0, & x' &\neq x, \\ \int \delta(x - x') \psi(x) dx &= \psi(x'). \end{aligned} \quad (16.43)$$

Si usamos $\delta(x - x')$ para la amplitud $\langle x | x' \rangle$ en la ecuación (16.38), dicha ecuación se satisface. Nuestro resultado es, entonces, que para nuestros estados base en x , la condición correspondiente a (16.36) es

$$\langle x' | x \rangle = \delta(x - x'). \quad (16.44)$$

Hemos completado ya las modificaciones necesarias de nuestras ecuaciones básicas necesarias para tratar con el continuo de estados base correspondientes a los puntos de una línea. La extensión a tres dimensiones es bastante clara; primero reemplazamos la coordenada x por el vector $\mathbf{r}$. Luego las integrales sobre x se reemplazan por integrales sobre x , y y z . En otras palabras, se convierten en integrales de volumen. Finalmente, la función δ unidimensional se debe reemplazar simplemente por el producto de tres funciones δ , una en x , una en y y la otra en z , $\delta(x - x')\delta(y - y')\delta(z - z')$. Reuniendo todo, obtenemos el siguiente conjunto de ecuaciones para las amplitudes para una partícula en tres dimensiones:

$$\langle \phi | \psi \rangle = \int \langle \phi | \mathbf{r} \rangle \langle \mathbf{r} | \psi \rangle dV, \quad (16.45)$$

$$\begin{aligned} \langle \mathbf{r} | \psi \rangle &= \psi(\mathbf{r}), \\ \langle \mathbf{r} | \phi \rangle &= \phi(\mathbf{r}), \end{aligned} \quad (16.46)$$

$$\langle \phi | \psi \rangle = \int \phi^*(\mathbf{r}) \psi(\mathbf{r}) dV, \quad (16.47)$$

$$\langle \mathbf{r}' | \mathbf{r} \rangle = \delta(x - x') \delta(y - y') \delta(z - z'). \quad (16.48)$$

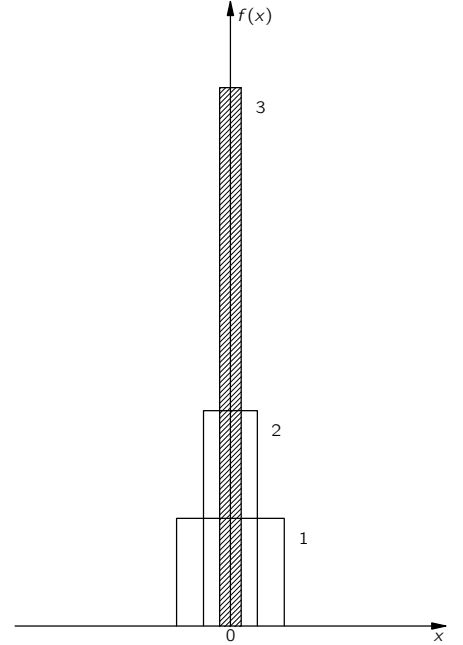


Figura 16-2. Un conjunto de funciones, todas de área uno, que se parecen más y más a $\delta(x)$.

¿Qué ocurre cuando hay más de una partícula? Les hablaremos de cómo manejar dos partículas, y ustedes verán fácilmente lo que deben hacer si quieren tratar un número más grande. Supongan que hay dos partículas, que podemos llamar partícula num. 1 y partícula num. 2. ¿Qué usaremos como estados base? Se puede describir un conjunto perfectamente adecuado diciendo que la partícula 1 está en x_1 y la partícula 2 en x_2 , lo cual podemos escribir en la forma $|x_1, x_2\rangle$. Observen que describir la posición de *una sola partícula no define* un estado base. Cada estado base debe definir la condición en que se encuentra todo el sistema. No deben pensar que cada partícula se mueve independientemente como una onda en tres dimensiones. Cualquier estado físico $|\psi\rangle$ se puede definir dando todas las amplitudes $\langle x_1, x_2 | \psi \rangle$ para encontrar a las dos partículas en x_1 y x_2 . Esta amplitud generalizada es, por lo tanto, una función de *dos* conjuntos de coordenadas x_1 y x_2 . Como pueden ver, esta función no es una onda, en el sentido de una oscilación que se mueve en tres dimensiones. En general, tampoco es sólo un producto de dos ondas individuales, una para cada partícula. Generalmente, es una especie de onda en las seis dimensiones definidas por x_1 y x_2 . Si en la naturaleza hay dos partículas que interactúan, no hay manera de describir lo que le ocurre a una de las partículas, tratando de escribir una función de onda para ella sola. Las famosas paradojas que consideramos en capítulos anteriores —en las cuales se sostenía que las mediciones hechas sobre una partícula podían decir lo que le iba a suceder a otra partícula, o podían destruir una interferencia— han dado muchos dolores de cabeza a la gente, porque han tratado de pensar en la función de onda de una partícula sola, en vez de la función de onda correcta en las coordenadas de ambas partículas. Sólo se puede dar la descripción completa de manera correcta en términos de funciones de las coordenadas de ambas partículas.

16-5 La ecuación de Schrödinger

Hasta ahora sólo nos hemos ocupado de cómo podemos describir estados que implican un electrón que se encuentra en cualquier parte del espacio. Ahora tenemos que preocuparnos de incluir en nuestra descripción la física de lo que puede ocurrir en diversas circunstancias. Como antes, tenemos que preocuparnos de cómo varían los estados en el tiempo. Si tenemos un estado $|\psi\rangle$, el cual pasa a otro estado $|\psi'\rangle$ un poco después, podemos describir la situación para todo tiempo haciendo que la función de onda —que es simplemente la amplitud $\langle \mathbf{r} | \psi \rangle$ — sea función del tiempo y de la posición. Una partícula en una situación dada puede ser descrita dando una función de onda que varía en el tiempo $\psi(\mathbf{r}, t) = \psi(x, y, z, t)$. Esta función de onda que varía en el tiempo describe la evolución de los estados sucesivos que se presentan a medida que transcurre el tiempo. La llamada “representación de coordenadas” —que da las proyecciones del estado $|\psi\rangle$ sobre los estados base $|\mathbf{r}\rangle$ — puede que no siempre sea la más conveniente, pero la consideraremos primero.

En el capítulo 8 estudiamos cómo varían los estados en el tiempo en términos del hamiltoniano H_{ij} . Vimos que la variación temporal de las diversas amplitudes estaba dada en términos de la ecuación matricial

$$i\hbar \frac{dC_i}{dt} = \sum_j H_{ij} C_j. \quad (16.49)$$

Esta ecuación indica que la variación temporal de cada amplitud C_i es proporcional a todas las demás amplitudes C_j con coeficientes H_{ij} .

¿Cuál esperaríamos que fuera el aspecto de la ecuación (16.49) al usar los estados base $|x\rangle$ del medio continuo? Recordemos primero que también es posible escribir la ecuación (16.49) en la forma

$$i\hbar \frac{d}{dt} \langle i | \psi \rangle = \sum_j \langle i | \hat{H} | j \rangle \langle j | \psi \rangle.$$

Ahora está claro qué es lo que debemos hacer. Para la representación x debemos esperar

$$i\hbar \frac{\partial}{\partial t} \langle x | \psi \rangle = \int \langle x | \hat{H} | x' \rangle \langle x' | \psi \rangle dx'. \quad (16.50)$$

La suma sobre los estados base $|j\rangle$, es remplazada por una integral sobre x' . Como $\langle x | \hat{H} | x' \rangle$ debe ser cierta función de x y x' , podemos escribirla como $H(x, x')$ —que corresponde al H_{ij} de la ecuación (16.49)—. Luego, la ecuación (16.50) es igual a

$$i\hbar \frac{\partial}{\partial t} \psi(x) = \int H(x, x') \psi(x') dx' \quad (16.51)$$

con

$$H(x, x') \equiv \langle x | \hat{H} | x' \rangle.$$

De acuerdo con la ecuación (16.51), la rapidez de cambio de ψ en x dependería del valor de ψ en todos los otros puntos x' ; el factor $H(x, x')$ es la amplitud por unidad de tiempo para que el electrón salte de x' a x . *Sin embargo, en la naturaleza ocurre que esta amplitud es cero, excepto para puntos x' muy cercanos a x .* Esto significa —como vimos en el ejemplo de la cadena de átomos al principio del capítulo, ecuación (16.12)— que el lado derecho de la ecuación (16.51) se puede expresar completamente en función de ψ y de las derivadas de ψ respecto a x , todas evaluadas en el punto x .

Para una partícula que se mueve libremente en el espacio sin ninguna fuerza, ninguna perturbación, la ley física correcta es

$$\int H(x, x') \psi(x') dx' = -\frac{\hbar^2}{2m} \frac{\partial^2}{\partial x^2} \psi(x).$$

¿De dónde la obtuvimos? De ninguna parte. No es posible deducirla de nada que conozcan. Emergió de la mente de Schrödinger, en su lucha por conseguir una comprensión de las observaciones experimentales del mundo real. Quizá tengan una indicación de por qué debe ser así, si piensan nuestra derivación de la ecuación (16.12), que se obtuvo examinando la propagación de un electrón en un cristal.

Por supuesto, las partículas libres no son muy interesantes. ¿Qué ocurre si aplicamos fuerzas a la partícula? Bueno, si la fuerza de una partícula se puede describir en términos de un potencial escalar $V(x)$ —lo cual significa que estamos pensando en fuerzas eléctricas, y no en fuerzas magnéticas— y si nos limitamos a energías bajas, de modo que podamos ignorar las complicaciones de los movimientos relativistas, entonces el hamiltoniano que se ajusta al mundo real es

$$\int H(x, x') \psi(x') dx' = -\frac{\hbar^2}{2m} \frac{\partial^2}{\partial x^2} \psi(x) + V(x) \psi(x). \quad (16.52)$$

Una vez más, tienen una indicación del origen de esta ecuación si regresan al movimiento de un electrón en un cristal, y ven cómo se tendrían que modificar las ecuaciones si la energía del electrón variara lentamente de una posición atómica a otra —como podría ocurrir si hubiera un campo eléctrico aplicado al cristal—. Entonces el término E_0 de la ecuación (16.7) variaría lentamente con la posición, y correspondería al nuevo término que hemos agregado en la ecuación (16.52).

[Puede que se estén preguntando por qué pasamos directamente de la ecuación (16.51) a la ecuación (16.52) en vez de darles simplemente la función correcta para la amplitud $H(x, x') = \langle x | \hat{H} | x' \rangle$. Lo hicimos porque sólo se puede escribir $H(x, x')$ en términos de funciones algebraicas extrañas, aunque toda la integral del lado derecho de la ecuación (16.51) sale en términos de cosas a las cuales están acostumbrados. Si realmente tienen curiosidad, se puede escribir $H(x, x')$ en la siguiente forma:

$$H(x, x') = -\frac{\hbar^2}{2m} \delta''(x - x') + V(x) \delta(x - x'),$$

donde δ'' significa la segunda derivada de la función delta. Esta función bastante extraña se puede remplazar por un operador diferencial algebraico más conveniente, que es completamente equivalente:

$$H(x, x') = \left\{ -\frac{\hbar^2}{2m} \frac{\partial^2}{\partial x^2} + V(x) \right\} \delta(x - x').$$

No usaremos estas expresiones, sino que trabajaremos directamente con la forma en la ecuación (16.52).]

Si ahora usamos la expresión que tenemos en (16.52) para la integral en (16.50), obtenemos la siguiente ecuación diferencial para $\psi(x) = \langle x | \psi \rangle$:

$$i\hbar \frac{\partial \psi}{\partial t} = -\frac{\hbar^2}{2m} \frac{\partial^2}{\partial x^2} \psi(x) + V(x)\psi(x). \quad (16.53)$$

Es bastante evidente lo que tendríamos que usar en vez de la ecuación (16.53) si estuviésemos interesados en el movimiento en tres dimensiones. Los únicos cambios consisten en remplazar $\partial^2/\partial x^2$ por

$$\nabla^2 = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2},$$

y $V(x)$ se remplaza con $V(x, y, z)$. La amplitud $\psi(x, y, z)$ para un electrón que se mueve en un potencial $V(x, y, z)$ obedece la ecuación diferencial

$$i\hbar \frac{\partial \psi}{\partial t} = -\frac{\hbar^2}{2m} \nabla^2 \psi + V\psi. \quad (16.54)$$

Se llama ecuación de Schrödinger, y fue la primera ecuación de la mecánica cuántica que se conoció. Fue propuesta por Schrödinger antes de que se descubriese cualquiera de las ecuaciones cuánticas que hemos descrito en este libro.

Aunque hemos enfocado el tema de una manera completamente diferente, el gran momento histórico que marcó el nacimiento de la descripción cuántica de la materia, fue cuando Schrödinger propuso su ecuación en 1926. En el curso de muchos años, la estructura atómica interna de la materia había sido un gran misterio. Nadie había podido comprender qué era lo que mantenía unida la materia, por qué había enlaces químicos y, especialmente, cómo podía ser que los átomos pudieran ser estables. Aunque Bohr había podido dar una descripción de los movimientos internos de un electrón en un átomo de hidrógeno, la cual parecía explicar el espectro observado de la luz emitida por este átomo, la razón por la que los electrones se movían de este modo seguía siendo un misterio. El descubrimiento que hizo Schrödinger de las ecuaciones de movimiento apropiadas para los electrones en escala atómica, suministró una teoría a partir de la cual se podían calcular los fenómenos atómicos cuantitativamente, en forma precisa y en detalle. En principio, la ecuación de Schrödinger es capaz de explicar todos los fenómenos atómicos, excepto aquellos en que interviene el magnetismo y la relatividad. Explica los niveles de energía de un átomo y todo lo referente al enlace químico. Sin embargo, esto es verdad sólo en principio —las matemáticas pronto se hacen demasiado complicadas como para resolver exactamente sólo los problemas más simples—. Únicamente los átomos de hidrógeno y de helio han sido calculados con gran precisión. Sin embargo, con diversas aproximaciones, algunas bastante burdas, se pueden comprender muchos aspectos más complicados referentes a los átomos y a los enlaces químicos. Les hemos mostrado algunas de estas aproximaciones en los capítulos anteriores.

Tal como la hemos escrito, la ecuación de Schrödinger no toma en cuenta ningún efecto magnético. Es posible tomar en cuenta estos efectos, en forma aproximada, agregando algunos términos más a la ecuación. No obstante, como lo vimos en el volumen II, el magnetismo es esencialmente un efecto relativista, por lo que una descripción correcta del movimiento de un electrón en un campo electromagnético arbitrario, sólo se puede discutir con una ecuación relativista apropiada. Dirac descubrió la ecuación relativista correcta para el movimiento de un electrón, un año después que Schrödinger propuso su ecuación, y dicha ecuación asume una forma completamente diferente. No seremos capaces de discutirla aquí.

Antes de continuar y examinar algunas consecuencias de la ecuación de Schrödinger, queremos mostrarles el aspecto que tiene para un sistema con gran número de partículas. No usaremos para nada la ecuación, sólo queremos mostrársela para recalcar que la función de onda ψ no es simplemente una onda

ordinaria en el espacio, sino una función de muchas variables. Si hay muchas partículas, la ecuación se convierte en

$$i\hbar \frac{\partial \psi(\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3, \dots)}{\partial t} = \sum_i -\frac{\hbar^2}{2m_i} \left\{ \frac{\partial^2 \psi}{\partial x_i^2} + \frac{\partial^2 \psi}{\partial y_i^2} + \frac{\partial^2 \psi}{\partial z_i^2} \right\} + V(\mathbf{r}_1, \mathbf{r}_2, \dots) \psi. \quad (16.55)$$

La función potencial V es lo que corresponde clásicamente a la energía potencial total de todas las partículas. Si no hay fuerzas externas actuando sobre las partículas, la función V es simplemente la energía electrostática de interacción de todas las partículas. Es decir, que si la i -ésima partícula tiene carga $Z_i q_e$, entonces la función V es simplemente†

$$V(\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3, \dots) = \sum_{\substack{\text{todos} \\ \text{los pares}}} \frac{Z_i Z_j}{r_{ij}} e^2. \quad (16.56)$$

16-6 Niveles de energía cuantizados

En un capítulo posterior examinaremos en detalle una solución de la ecuación de Schrödinger para un ejemplo particular. Sin embargo, ahora queremos mostrarles cómo surge una de las consecuencias más notables de la ecuación de Schrödinger —el hecho sorprendente de una ecuación diferencial, en la que únicamente intervienen funciones continuas de variables continuas en el espacio, puede dar lugar a efectos cuánticos tales como los niveles discretos de energía de un átomo—. El hecho esencial es comprender cómo puede ser que un electrón confinado en una región del espacio por cierto tipo de “pozo” de potencial, debe tener necesariamente sólo una u otra energía de un conjunto bien definido de energías discretas.

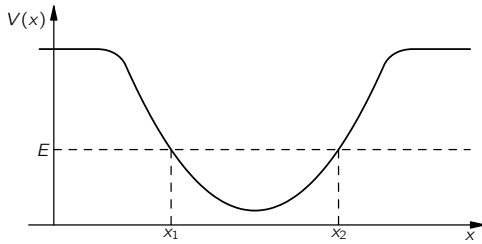


Figura 16-3. Un pozo de potencial para una partícula que se mueve en la dirección x .

Supongan que consideramos a un electrón en una situación unidimensional, en la que su energía potencial varía con x de la manera descrita por la gráfica de la figura 16-3. Supondremos que este potencial es estático —no varía con el tiempo—. Tal como lo hemos hecho muchas veces antes, queremos buscar soluciones correspondientes a estados de energía definida, lo cual significa, de frecuencia definida. Intentemos una solución de la forma

$$\psi = a(x)e^{-iEt/\hbar}. \quad (16.57)$$

Si sustituimos esta función en la ecuación de Schrödinger, encontramos que la función $a(x)$ debe satisfacer la siguiente ecuación diferencial:

$$\frac{d^2 a(x)}{dx^2} = \frac{2m}{\hbar^2} [V(x) - E] a(x). \quad (16.58)$$

Esta ecuación dice que en cada x la segunda derivada de $a(x)$ respecto a x es proporcional a $a(x)$, el coeficiente de proporcionalidad está dado por la cantidad $(2m/\hbar^2)(V - E)$. La segunda derivada de $a(x)$ es la rapidez de cambio de su pendiente. Si el potencial V es mayor que la energía E de la partícula, la rapidez de cambio de la pendiente de $a(x)$ tendrá el mismo signo que $a(x)$. Esto significa que la curva de $a(x)$ será cóncava respecto al eje x . Es decir, tendrá, más o menos, el comportamiento de la función exponencial positiva o negativa, $e^{\pm x}$. Esto significa que en la región a la izquierda de x_1 , en la figura 16-3, donde V es mayor que la energía E supuesta, la función $a(x)$ tendrá el aspecto de alguna de las curvas mostradas en la parte (a) de la figura 16-4.

† Estamos usando la convención de los volúmenes anteriores, según la cual $e^2 \equiv q_e^2/4\pi\epsilon_0$.

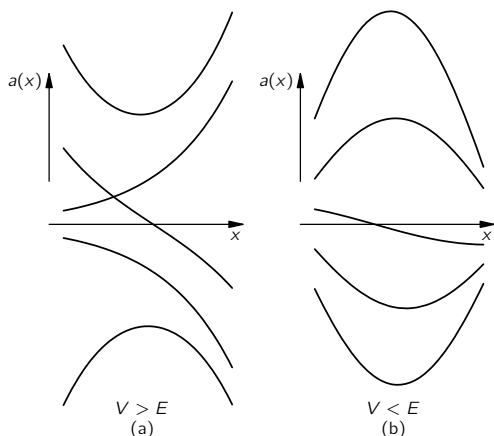


Figura 16-4. Formas posibles de la función de onda $a(x)$ para $V > E$ y para $V < E$.

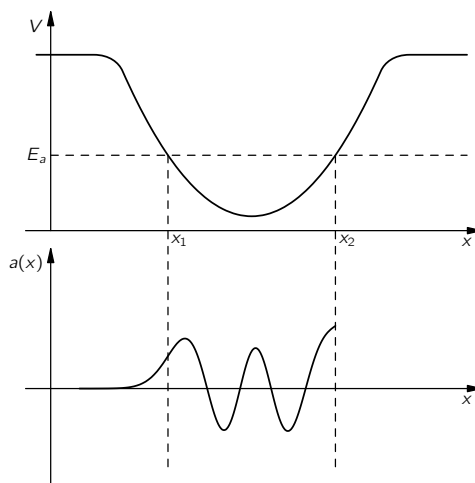


Figura 16-5. Una función de onda correspondiente a la energía E_a que va a cero para x negativa.

Por otra parte, si la función potencial V es menor que la energía E , la segunda derivada de $a(x)$ respecto a x tiene signo opuesto al de $a(x)$, y la curva de $a(x)$ siempre será cóncava hacia el eje x , como alguna de las curvas mostradas en la parte (b) de la figura 16-4. En esa región, la solución tiene, en cada trozo, aproximadamente la forma de una curva sinusoidal.

Ahora veamos si podemos construir gráficamente una solución para la función $a(x)$ que corresponda a una partícula de energía E_a en el potencial V que muestra la figura 16-3. Dado que estamos tratando de describir una situación en que la partícula está unida *dentro* del pozo de potencial, queremos buscar soluciones en las que la amplitud de onda tome valores muy pequeños cuando x se aleje del pozo de potencial. Podemos imaginar fácilmente una curva como la que muestra la figura 16-5, la cual tiende a cero para valores negativos grandes de x y crece suavemente al aproximarse a x_1 . Como V es igual a E_a en x_1 , la curvatura de la función se hace cero en este punto. Entre x_1 y x_2 , la cantidad $V - E_a$ siempre es un número negativo, por lo que la función $a(x)$ siempre es cóncava hacia el eje, y la curvatura es tanto mayor, mientras más grande es la diferencia entre E_a y V . Si continuamos la curva dentro de la región entre x_1 y x_2 , será más o menos como muestra la figura 16-5.

Ahora continuemos esta curva en la región a la derecha de x_2 . Ahí se aleja del eje y va a valores positivos grandes, como se muestra en la figura 16-6. Para la energía E_a que hemos elegido, la solución $a(x)$ crece continuamente al aumentar x . En realidad, su *curvatura* también crece (si el potencial sigue siendo constante). La amplitud crece rápidamente a valores inmensos. ¿Qué significa esto? Significa simplemente que la partícula no está “unida” en el pozo de potencial. Es infinitamente más probable que se encuentre *fuera* que *dentro* del pozo. Para la solución que hemos desarrollado, es más probable que el electrón se encuentre en $x = +\infty$ que en cualquier otra parte. Hemos fracasado en encontrar una solución para una partícula unida.

Intentemos con otra energía, digamos una un poco mayor que E_a —por ejemplo, la energía E_b de la figura 16-7—. Si empezamos con la misma condición a la izquierda, obtenemos la solución mostrada en la mitad inferior de la figura 16-7. Al principio parecía que iba a ser mejor, pero termina siendo tan mala como la solución para E_a —excepto que, ahora, $a(x)$ se hace más y más *negativa* a medida que avanzamos hacia los valores grandes de x .

Puede que ésa sea la clave. Como cambiar la energía un poco de E_a a E_b causa que la curva salte de un lado del eje al otro, quizá haya alguna energía

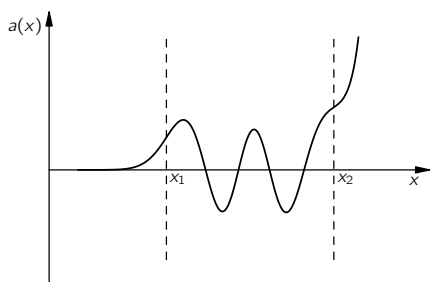


Figura 16-6. La función de onda $a(x)$ de la figura 16-5 continuada más allá de x_2 .

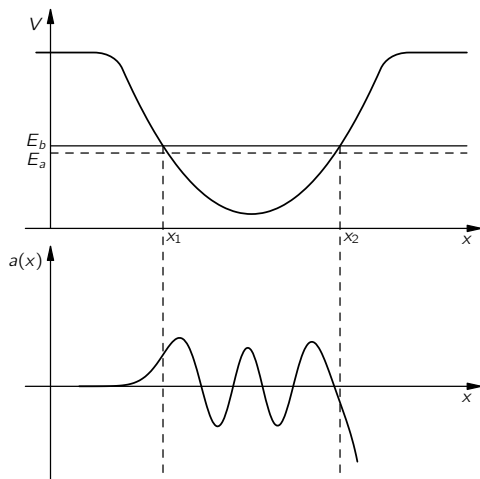


Figura 16-7. La función de onda $a(x)$ para una energía E_b mayor que E_a .

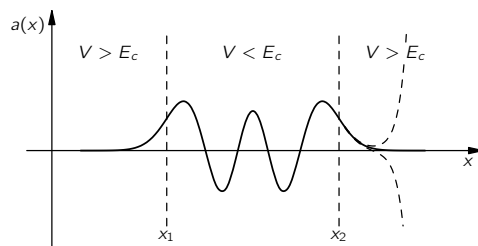


Figura 16-8. Una función de onda para la energía E_c entre E_a y E_b .

entre E_a y E_b para la cual la curva tienda a cero para grandes valores de x . La hay, en realidad, y hemos esbozado cuál podría ser el aspecto de la solución en la figura 16-8.

Se darán cuenta que la solución que hemos dibujado en la figura es muy especial. Si aumentásemos o disminuyésemos la energía, aunque fuese lo más mínimo, la función tomaría alguna forma como la de las curvas discontinuas mostradas en la figura 16-8, y no tendríamos las condiciones apropiadas para una partícula unida. Hemos obtenido el resultado de que si una partícula tiene que estar unida en un pozo de potencial, sólo lo puede hacer si tiene una energía muy definida.

¿Esto significa que sólo hay una energía para una partícula unida en un pozo de potencial? No. Hay otras energías posibles, pero no energías demasiado cercanas a E_c . Observen que la función de onda que hemos dibujado en la figura 16-8 cruza el eje cuatro veces en la región entre x_1 y x_2 . Si eligiéramos una energía bastante más baja que E_c , podríamos tener una solución que cruce el eje sólo tres veces, sólo dos veces, sólo una vez, o ninguna. Las soluciones posibles están dibujadas esquemáticamente en la figura 16-9. (También puede haber otras soluciones correspondientes a valores de la energía más altos que los indicados.) Nuestra conclusión es que si una partícula está unida en un pozo de potencial, su energía puede tomar únicamente ciertos valores especiales de un espectro discreto de energía. Ven cómo una ecuación diferencial puede describir el hecho básico de la física cuántica.

Agreguemos algo más. Si la energía E está por encima del máximo del pozo de potencial, ya no hay soluciones discretas y cualquier energía está permitida. Esas soluciones corresponden a la dispersión de partículas libres por un pozo de potencial. Hemos visto un ejemplo de esas soluciones cuando consideramos los efectos de átomos de impureza en un cristal.

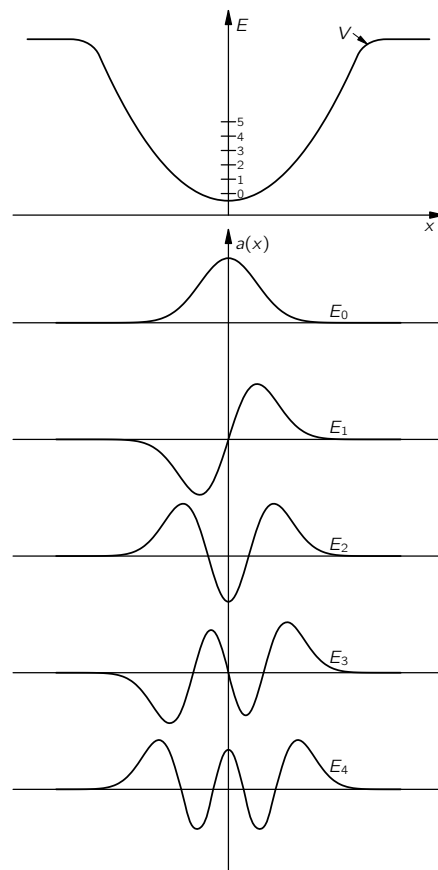


Figura 16-9. La función $a(x)$ para los cinco estados unidos de energía más bajos.

Simetría y leyes de conservación

17-1 Simetría

En física clásica hay un cierto número de cantidades que se *conservan* —tales como el momentum, la energía y el momentum angular—. También existen teoremas de conservación de las cantidades correspondientes en la mecánica cuántica. Lo más maravilloso de la mecánica cuántica es que los teoremas de conservación se pueden derivar, en cierto sentido, de algo distinto, mientras que en la mecánica clásica son prácticamente los puntos de partida de las leyes. (En la mecánica clásica hay maneras de hacer algo análogo a lo que haremos en la mecánica cuántica, pero solamente se puede hacer en un nivel muy avanzado.) En la mecánica cuántica, sin embargo, las leyes de conservación están profundamente relacionadas con el principio de superposición de amplitudes y con la simetría de sistemas físicos sometidos a diversos cambios. Éste es el tema del presente capítulo. Aunque aplicaremos estas ideas principalmente a la conservación del momentum angular, el punto esencial es que los teoremas de conservación de cualquier cantidad están —en la mecánica cuántica— relacionados con las simetrías del sistema.

Comenzamos, por lo tanto, estudiando la cuestión de las simetrías de sistemas. Un ejemplo muy simple es el ion de hidrógeno molecular —nos daría igual tomar la molécula de amoníaco— en el que hay dos estados. Para el ion de hidrógeno molecular, tomamos como estados base, uno en el cual el electrón estaba colocado cerca del protón número 1 y otro en el cual el electrón estaba cerca del protón número 2. Los dos estados —a los que llamamos $|1\rangle$ y $|2\rangle$ — se muestran en la figura 17-1(a). Ahora bien, mientras los dos núcleos sean exactamente iguales, hay una cierta *simetría* en este sistema físico. Es decir, si *reflejásemos* al sistema en el plano a la mitad de los dos protones —con lo cual queremos dar a entender que todo lo que hay de un lado del plano, pasa a la posición simétrica del otro lado— obtendríamos la situación de la figura 17-1(b). Como los protones son idénticos, la *operación de reflexión* convierte a $|1\rangle$ en $|2\rangle$ y a $|2\rangle$ en $|1\rangle$. Llamaremos esta operación de reflexión $\hat{P}$ y escribiremos

$$\hat{P}|1\rangle = |2\rangle, \quad \hat{P}|2\rangle = |1\rangle. \quad (17.1)$$

Por lo que nuestra $\hat{P}$ es un operador, en el sentido de que “hace algo” a un estado para convertirlo en un nuevo estado. Lo interesante es que $\hat{P}$ operando sobre *cualquier* estado produce algún *otro* estado del sistema.

Además $\hat{P}$, como cualquiera de los operadores que hemos descrito, tiene elementos matriciales, los cuales se pueden definir mediante la notación sencilla de costumbre. Es decir,

$$P_{11} = \langle 1 | \hat{P} | 1 \rangle \quad \text{y} \quad P_{12} = \langle 1 | \hat{P} | 2 \rangle$$

son los elementos matriciales que obtenemos si multiplicamos a $\hat{P}|1\rangle$ y a $\hat{P}|2\rangle$ en la izquierda por $\langle 1|$. Según la ecuación (17.1), éstos son

$$\begin{aligned} \langle 1 | \hat{P} | 1 \rangle &= P_{11} = \langle 1 | 2 \rangle = 0, \\ \langle 1 | \hat{P} | 2 \rangle &= P_{12} = \langle 1 | 1 \rangle = 1. \end{aligned} \quad (17.2)$$

Del mismo modo podemos obtener P_{21} y P_{22} . La matriz de $\hat{P}$ —respecto al sistema de base $|1\rangle$ y $|2\rangle$ — es

$$P = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}. \quad (17.3)$$

17-1 Simetría

17-2 Simetría y conservación

17-3 Las leyes de conservación

17-4 Luz polarizada

17-5 La desintegración de la Λ^0

17-6 Resumen de las matrices de rotación

Revisar

Volumen I, capítulo 52,
Simetría en las leyes físicas.

Referencia:

A. R. Edmonds, *Angular Momentum in Quantum Mechanics*, Princeton University Press, 1957.

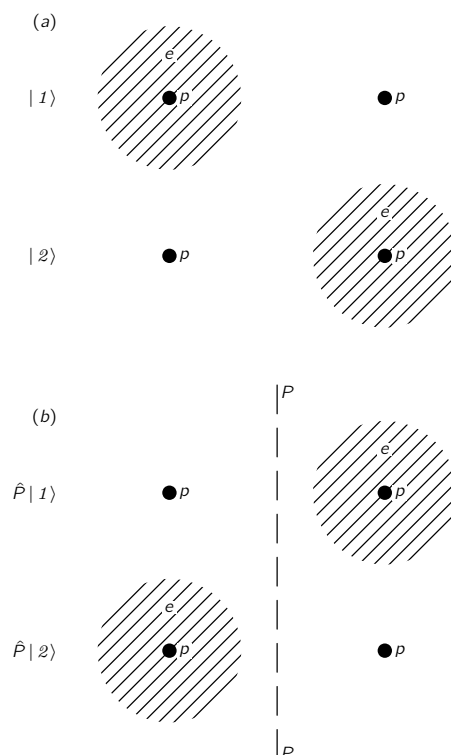


Figura 17-1. Si los estados $|1\rangle$ y $|2\rangle$ se reflejan en el plano P - P , se convierten en $|2\rangle$ y $|1\rangle$, respectivamente.

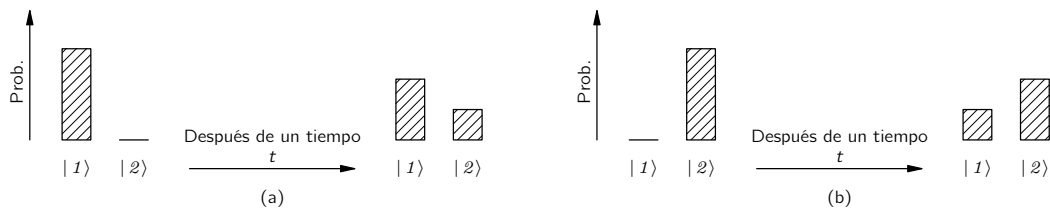


Figura 17-2. En un sistema simétrico, si un estado puro $|1\rangle$ se transforma como muestra la parte (a), un estado puro $|2\rangle$ se transformará como en la parte (b).

Vemos una vez más que las palabras *operador* y *matriz*, en la mecánica cuántica, son prácticamente intercambiables. Hay pequeñas diferencias técnicas —como la diferencia entre “numeral” y “número”— pero la distinción es un poco pedante, por lo que no tenemos que preocuparnos. Por lo tanto, ya sea que $\hat{P}$ defina una operación, o que se utilice para definir una matriz de números, lo llamaremos indistintamente operador o matriz.

Ahora nos gustaría señalar algo. *Supondremos* que la *física* del sistema total del ion de hidrógeno molecular es *simétrica*. No tiene que serlo —depende, por ejemplo, de lo que haya cerca de éste—. Pero si el sistema es simétrico, la siguiente idea sería verdadera. Supongan que empezamos a $t = 0$ con el sistema en el estado $|1\rangle$ y encontramos que, después de un intervalo de tiempo t , el sistema está en un estado más complicado —en alguna combinación lineal de los dos estados base—. Recuerden que en el capítulo 8 representamos a “transcurrir un intervalo de tiempo” mediante la multiplicación por el operador $\hat{U}$. Esto significa que el sistema un momento después —por ejemplo, 15 segundos— estaría en algún otro estado. Por ejemplo, podría ser $\sqrt{2/3}$ partes del estado $|1\rangle$ y $i\sqrt{1/3}$ partes del estado $|2\rangle$, y escribiríamos

$$|\psi \text{ en } 15 \text{ s}\rangle = \hat{U}(15, 0) |1\rangle = \sqrt{2/3} |1\rangle + i\sqrt{1/3} |2\rangle. \quad (17.4)$$

Ahora nos preguntamos, ¿qué sucede si comenzamos con el sistema en el estado *simétrico* $|2\rangle$ y esperamos 15 segundos bajo las *mismas condiciones*? Es claro que si el mundo es simétrico —como estamos suponiendo— obtendríamos el estado simétrico de (17.4):

$$|\psi \text{ en } 15 \text{ s}\rangle = \hat{U}(15, 0) |2\rangle = \sqrt{2/3} |2\rangle + i\sqrt{1/3} |1\rangle. \quad (17.5)$$

Estas ideas se presentan en la figura 17-2. Así, si la *física* de un sistema es simétrica respecto a algún plano y calculamos el comportamiento de un estado particular, también conocemos el comportamiento del estado que obtendríamos reflejando el estado original en el plano de simetría.

Nos gustaría decir lo mismo, pero de un modo un poco más general —lo que significa un poco más abstracto—. Sea $\hat{Q}$ una de las operaciones cualesquiera que ustedes pueden realizar sobre un sistema *sin cambiar la física*. Por ejemplo, $\hat{Q}$ podría ser muy bien $\hat{P}$, operación de *reflexión* en el plano entre los dos átomos de una molécula de hidrógeno. O, en un sistema con dos electrones, podría ser la operación de *intercambiar* los dos electrones. Otra posibilidad podría ser, en un sistema con simetría esférica, la operación de una *rotación* de todo el sistema, en un ángulo determinado, alrededor de un eje —lo cual no cambiaría la física—. Naturalmente, queremos dar a $\hat{Q}$ una notación especial para cada caso particular. En específico, definiremos normalmente $\hat{R}_y(\theta)$ como la operación de “rotar al sistema alrededor del eje y en el ángulo θ ”. Por $\hat{Q}$ indicamos cualquiera de los operadores que hemos descrito o cualquier otro —que no cambie la situación física básica—.

Consideremos algunos otros ejemplos. Si tenemos un átomo sin *ningún* campo *magnético externo* o sin *ningún* campo *eléctrico externo*, yuviésemos que rotar las coordenadas alrededor de cualquier eje, sería el mismo sistema físico. De nuevo, la molécula de amoníaco es simétrica respecto a una reflexión en un

plano paralelo al de los tres hidrógenos —siempre que no haya ningún campo eléctrico—. Si hay un campo eléctrico, cuando hacemos una reflexión, tendríamos que cambiar también el campo eléctrico, lo que cambia el problema físico. Pero si no tenemos campo externo, la molécula es simétrica.

Ahora consideremos un caso general. Supongan que empezamos con el estado $|\psi_1\rangle$, y que después de algún tiempo, en determinadas condiciones físicas, ha pasado a ser el estado $|\psi_2\rangle$. Podemos escribir

$$|\psi_2\rangle = \hat{U} |\psi_1\rangle. \quad (17.6)$$

[Pueden pensar en la ecuación (17.4).] Ahora imaginen que llevamos a cabo la operación $\hat{Q}$ sobre todo el sistema. El estado $|\psi_1\rangle$ se transformará en el estado $|\psi'_1\rangle$, que podemos escribir también como $\hat{Q}|\psi_1\rangle$. El estado $|\psi_2\rangle$ también se cambiará al estado $|\psi'_2\rangle = \hat{Q}|\psi_2\rangle$. Ahora bien, *si la física* es simétrica bajo $\hat{Q}$ (no olviden el *si*; no es una propiedad general de los sistemas), entonces, esperando el mismo tiempo en las mismas condiciones, deberíamos tener

$$|\psi'_2\rangle = \hat{U} |\psi'_1\rangle. \quad (17.7)$$

[Como la ecuación (17.5).] Pero podemos escribir $\hat{Q}|\psi_1\rangle$ para $|\psi'_1\rangle$ y $\hat{Q}|\psi_2\rangle$ para $|\psi'_2\rangle$, por lo que (17.7) se puede también escribir como

$$\hat{Q}|\psi_2\rangle = \hat{U} \hat{Q} |\psi_1\rangle. \quad (17.8)$$

Si ahora remplazamos a $|\psi_2\rangle$ por $\hat{U}|\psi_1\rangle$ —ecuación (17.6)— obtenemos

$$\hat{Q} \hat{U} |\psi_1\rangle = \hat{U} \hat{Q} |\psi_1\rangle. \quad (17.9)$$

No es difícil entender lo que esto significa. Respecto al ion de hidrógeno, nos dice que “hacer una reflexión y esperar un rato” —la expresión en el lado derecho de (17.9)— es lo mismo que “esperar un rato y luego hacer una reflexión” —la expresión en el lado izquierdo de (17.9)—. Éstas deberían ser iguales, siempre que U no varíe en una reflexión.

Como (17.9) es válida para *cualquier* estado inicial $|\psi_1\rangle$, realmente es una ecuación entre operadores:

$$\hat{Q} \hat{U} = \hat{U} \hat{Q}. \quad (17.10)$$

Esto es lo que queríamos obtener —*un enunciado matemático de simetría*—. Cuando la ecuación (17.10) se cumple, decimos que los operadores $\hat{U}$ y $\hat{Q}$ *conmutan*. Podemos entonces *definir* “simetría” como sigue: un sistema físico es *simétrico* respecto a la operación $\hat{Q}$, cuando $\hat{Q}$ conmuta con $\hat{U}$, la operación de transcurso de tiempo. [En términos de matrices, el producto de dos operadores es equivalente a la matriz producto, de este modo la ecuación (17.10) también sirve para las matrices Q y U , para un sistema que es simétrico frente a la transformación Q .]

Incidentalmente, puesto que para un tiempo infinitesimal por ϵ tenemos que $\hat{U} = 1 - i\hat{H}\epsilon/\hbar$ —donde $\hat{H}$ es el hamiltoniano de costumbre (véase el capítulo 8)— pueden ver que si (17.10) es válida, también lo es

$$\hat{Q} \hat{H} = \hat{H} \hat{Q}. \quad (17.11)$$

Por lo que (17.11) es el enunciado matemático de la condición de simetría de una situación física frente al operador $\hat{Q}$. Éste *define* una simetría.

17-2 Simetría y conservación

Antes de aplicar el resultado que acabamos de obtener, queremos discutir un poco más la idea de simetría. Supongan que tenemos una situación muy especial: después de que aplicamos $\hat{Q}$ a un estado, obtenemos el mismo estado. Éste es un caso muy especial, pero supongamos que ocurre así para un estado $|\psi_0\rangle$; esto es, que $|\psi'\rangle = \hat{Q}|\psi_0\rangle$ es físicamente el mismo estado que $|\psi_0\rangle$. Esto significa que

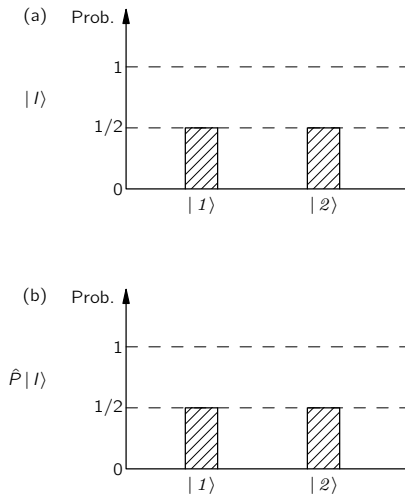


Figura 17-3. El estado $|I\rangle$ y el estado $\hat{P}|I\rangle$ que se obtiene reflejando $|I\rangle$ en el plano central.

$|\psi'\rangle$ es igual a $|\psi_0\rangle$, excepto por algún factor de fase.[†] ¿Cómo puede suceder esto? Por ejemplo, supongan que tenemos un ion H_2^+ en el estado que antes llamamos $|I\rangle$.[‡] Para este estado hay una amplitud igual para estar en los estados base $|1\rangle$ y $|2\rangle$. La figura 17-3(a) muestra las probabilidades en una gráfica de barras. Si aplicamos a $|I\rangle$ el operador de reflexión $\hat{P}$, éste hace que el estado se invierta, cambiando $|1\rangle$ por $|2\rangle$ y $|2\rangle$ por $|1\rangle$ —obtenemos las probabilidades que nos muestra la figura 17-3(b)—. Pero esto es exactamente el mismo estado $|I\rangle$. Si comenzamos con el estado $|II\rangle$, las probabilidades antes y después de la reflexión son iguales. Sin embargo, hay una diferencia si consideramos las *amplitudes*. Para el estado $|I\rangle$, las amplitudes son *iguales* después de la reflexión; pero para el estado $|II\rangle$, las amplitudes tienen signos opuestos. En otras palabras,

$$\begin{aligned}\hat{P}|I\rangle &= \hat{P}\left\{\frac{|1\rangle + |2\rangle}{\sqrt{2}}\right\} = \frac{|2\rangle + |1\rangle}{\sqrt{2}} = |I\rangle, \\ \hat{P}|II\rangle &= \hat{P}\left\{\frac{|1\rangle - |2\rangle}{\sqrt{2}}\right\} = \frac{|2\rangle - |1\rangle}{\sqrt{2}} = -|II\rangle.\end{aligned}\tag{17.12}$$

Si escribimos $\hat{P}|\psi_0\rangle = e^{i\delta}|\psi_0\rangle$, tenemos que $e^{i\delta} = 1$ para el estado $|I\rangle$ y $e^{i\delta} = -1$ para el estado $|II\rangle$.

Consideremos otro ejemplo. Supongan que tenemos un fotón polarizado RHC que se propaga en la dirección z . Si aplicamos una rotación alrededor del eje z , sabemos que esto sólo multiplica a la amplitud por $e^{i\phi}$, siendo ϕ el ángulo de rotación. Así, para la operación de rotación en este caso, δ es justamente igual al ángulo de rotación.

Ahora está claro que *si en realidad es verdad* que un operador $\hat{Q}$ sólo cambia la fase de un estado en cierto instante, digamos $t = 0$, *es verdad para siempre*. En otras palabras, si el estado $|\psi_1\rangle$ pasa al estado $|\psi_2\rangle$ después de un tiempo t , o sea

$$\hat{U}(t, 0)|\psi_1\rangle = |\psi_2\rangle\tag{17.13}$$

y si la simetría de la situación hace que

$$\hat{Q}|\psi_1\rangle = e^{i\delta}|\psi_1\rangle,\tag{17.14}$$

entonces también es verdad que

$$\hat{Q}|\psi_2\rangle = e^{i\delta}|\psi_2\rangle.\tag{17.15}$$

Esto es claro, puesto que

$$\hat{Q}|\psi_2\rangle = \hat{Q}\hat{U}|\psi_1\rangle = \hat{U}\hat{Q}|\psi_1\rangle,$$

y si $\hat{Q}|\psi_1\rangle = e^{i\delta}|\psi_1\rangle$, entonces

$$\hat{Q}|\psi_2\rangle = \hat{U}e^{i\delta}|\psi_1\rangle = e^{i\delta}\hat{U}|\psi_1\rangle = e^{i\delta}|\psi_2\rangle.$$

[La sucesión de igualdades se sigue de las ecuaciones (17.13) y (17.10) para un sistema simétrico, de (17.14), y del hecho de que un número como $e^{i\delta}$ conmuta con un operador.]

Así, con ciertas simetrías, algo que es verdad inicialmente, es verdad para cualquier tiempo. Pero ¿no es esto justamente una *ley de conservación*? ¡Sí! Afirma que si consideran al estado original y si, haciendo un pequeño cálculo, descubren que una operación, *que es una operación de simetría del sistema*, produce sólo una multiplicación por cierta fase, saben entonces que la misma propiedad será verdad en el estado final —la misma operación multiplica al estado

[†] A propósito, pueden demostrar que $\hat{Q}$ es necesariamente un *operador unitario* —lo que significa que si se aplica a $|\psi\rangle$, da algún número por $|\psi\rangle$, el número debe ser de la forma $e^{i\delta}$, donde δ es real—. Es algo sin importancia, y su demostración se basa en la siguiente observación. Cualquier operación como una reflexión o una rotación no pierde partículas, por lo que la normalización de $|\psi'\rangle$ y $|\psi\rangle$ debe ser la misma; pueden diferir solamente en un factor de fase puramente imaginario.

[‡] Véase la sección 10-1. Los estados $|I\rangle$ y $|II\rangle$ están invertidos en esta sección respecto a la discusión previa.

final por el mismo factor de fase—. Esto siempre es verdad, aunque pueda ser que no conozcamos nada más acerca del mecanismo interno del universo, el cual lleva a un sistema del estado inicial al final. Aunque no nos preocupemos en considerar los detalles de la maquinaria mediante la cual el sistema pasa de un estado a otro, aun así podemos decir que si algo está originalmente en un estado con un cierto carácter de simetría, y que si el hamiltoniano de ese algo es simétrico frente a esa operación de simetría, entonces el estado tendrá el mismo carácter de simetría para cualquier tiempo. Ésta es la base de todas las leyes de conservación de la mecánica cuántica.

Consideremos un ejemplo especial. Volvamos al operador $\hat{P}$. Primero queremos modificar un poco nuestra definición de $\hat{P}$. Queremos que $\hat{P}$ no sea exactamente una reflexión en el espejo, porque esto requiere definir el plano en el que colocamos el espejo. Hay una clase especial de reflexión que no requiere la especificación de un plano. Supongan que redefinimos la operación $\hat{P}$ así: primero, reflejan en un espejo situado en el plano z , de modo que z pasa a $-z$, x sigue siendo x y y sigue siendo y ; luego giran el sistema 180° alrededor del eje z , de modo que x pasa a $-x$ y y a $-y$. Todo esto se llama *inversión*. Cada punto se ha proyectado *a través del origen* a la posición diametralmente opuesta. Todas las coordenadas se han invertido. Seguiremos usando el símbolo $\hat{P}$ para esta operación. Esto se muestra en la figura 17-4. Es más fácil que una reflexión simple, porque no requiere que especifiquen qué sistema de coordenadas usan para la reflexión —sólo necesitan especificar el punto que es el centro de simetría—.

Supongamos ahora que tenemos un estado $|\psi_0\rangle$ que, sometido a la operación de inversión, pasa al $e^{i\delta}|\psi_0\rangle$; esto es,

$$|\psi'_0\rangle = \hat{P}|\psi_0\rangle = e^{i\delta}|\psi_0\rangle. \quad (17.16)$$

Supongan que volvemos a invertir. Después de *dos* inversiones nos encontramos precisamente donde partimos —nada ha cambiado—. Debemos tener

$$\hat{P}|\psi'_0\rangle = \hat{P}\hat{P}|\psi_0\rangle = |\psi_0\rangle.$$

Pero

$$\hat{P}\hat{P}|\psi_0\rangle = \hat{P}e^{i\delta}|\psi_0\rangle = e^{i\delta}\hat{P}|\psi_0\rangle = (e^{i\delta})^2|\psi_0\rangle.$$

De donde

$$(e^{i\delta})^2 = 1.$$

Entonces, *si el operador de inversión es una operación de simetría* de un estado, hay solamente dos posibilidades para $e^{i\delta}$:

$$e^{i\delta} = \pm 1,$$

lo que significa que

$$\hat{P}|\psi_0\rangle = |\psi_0\rangle \quad \text{o} \quad \hat{P}|\psi_0\rangle = -|\psi_0\rangle. \quad (17.17)$$

Clásicamente, si un estado es simétrico bajo una inversión, la operación reproduce el mismo estado. En la mecánica cuántica, sin embargo, hay las dos posibilidades: obtenemos el *mismo estado* o *menos* el mismo estado. Cuando obtenemos el *mismo* estado, $\hat{P}|\psi_0\rangle = |\psi_0\rangle$, decimos que el estado $|\psi_0\rangle$ tiene *paridad par*. Cuando se invierte el signo, de modo que $\hat{P}|\psi_0\rangle = -|\psi_0\rangle$, decimos que el estado tiene *paridad impar*. (El operador de inversión $\hat{P}$ se conoce también como operador de paridad.) El estado $|I\rangle$ del ion H_2^+ tiene paridad par; y el $|II\rangle$ tiene paridad impar —véase la ecuación (17.12)—. Hay, por supuesto, estados que no son simétricos frente a la operación $\hat{P}$; éstos son estados con paridad no definida. Por ejemplo, en el sistema H_2^+ , el estado $|I\rangle$ tiene paridad par, el estado $|II\rangle$ tiene paridad impar y el estado $|1\rangle$ no tiene paridad definida.

Cuando decimos que una operación como la inversión se aplica “en un sistema físico”, podemos considerarla de dos modos. Podemos pensar en *mover físicamente* todo lo que está en $\mathbf{r}$ al punto inverso $-\mathbf{r}$, o podemos pensar en *ver* al mismo

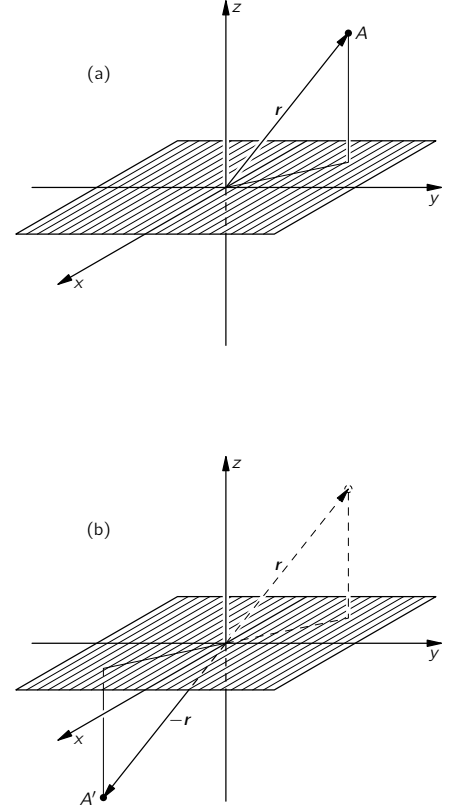


Figura 17-4. La operación de inversión, $\hat{P}$. Cualquier cosa que se encuentre en el punto A en (x, y, z) se traslada al punto A' en $(-x, -y, -z)$.

sistema desde un nuevo sistema de referencia x', y', z' , relacionado con el viejo mediante $x' = -x$, $y' = -y$ y $z' = -z$. Del mismo modo, cuando consideremos rotaciones, podemos pensar que el sistema ha sido girado, o que se han girado los ejes de coordenadas respecto a los cuales medimos el sistema, conservando al “sistema” fijo en el espacio. Generalmente, los dos puntos de vista son esencialmente equivalentes. Para una rotación son equivalentes, *excepto* que rotar a un *sistema* el ángulo θ , es como rotar el sistema de referencia el ángulo *menos* θ . En nuestras lecciones, por lo general hemos considerado lo que sucede cuando se hace una proyección sobre un nuevo conjunto de ejes. Lo que obtienen así es igual a lo que obtienen si dejan los ejes fijos y rotan el sistema *hacia atrás* en la misma cantidad. Al hacer eso, invierten los signos de los ángulos.†

Muchas *leyes* de la física —pero no todas— no varían por una reflexión o una inversión de coordenadas. Son *simétricas* respecto a una inversión. Las leyes de la electrodinámica, por ejemplo, no varían si cambiamos a x por $-x$, y por $-y$ y z por $-z$ en *todas* las ecuaciones. Así mismo con las leyes de gravedad y con las interacciones fuertes de física nuclear. Solamente las interacciones débiles —responsables de la desintegración β — no tienen esta simetría. (Discutimos esto detalladamente en el capítulo 52 del volumen I.) Por ahora no nos ocuparemos de la desintegración β . Entonces en cualquier sistema físico donde la desintegración β no produzca un efecto apreciable —un ejemplo sería la emisión de luz por un átomo— el hamiltoniano $\hat{H}$ y el operador $\hat{P}$ conmutarán. En estas circunstancias tenemos la siguiente proposición. Si un estado tiene originalmente paridad par, y si consideran la situación física en un instante posterior, aún tendrá paridad par. Por ejemplo, supongan que un átomo que está por emitir un fotón se encuentra en un estado que sabemos tiene paridad par. Si consideran todo —incluyendo al fotón— después de la emisión, aún tendrá paridad par (lo mismo ocurre si comienzan con paridad impar). Este principio se llama *conservación de la paridad*. Pueden ver por qué los conceptos “conservación de la paridad” y “simetría de reflexión” están tan íntimamente entrelazados en la mecánica cuántica. Aunque hasta hace unos pocos años se pensaba que la naturaleza siempre conservaba la paridad, ahora se sabe que esto *no* es verdad. Se ha descubierto que es falso porque la reacción de desintegración β no posee la simetría de inversión, la cual se encuentra en las otras leyes de la física.

Ahora podemos probar un teorema interesante (que es válido siempre que podamos despreciar las interacciones débiles): cualquier estado de energía definida no degenerado debe tener una paridad definida. Debe tener ya sea paridad par o paridad impar. (Recuerden que a veces hemos visto sistemas, en los cuales varios estados tienen la misma energía —decimos que tales estados son *degenerados*—. Nuestro teorema no sirve para ellos.)

Para un estado $|\psi_0\rangle$ de energía definida, sabemos que

$$\hat{H}|\psi_0\rangle = E|\psi_0\rangle, \quad (17.18)$$

donde E es sólo un número —la energía del estado—. Si tenemos *cualquier* operador $\hat{Q}$, el cual sea un operador simétrico del sistema, podemos probar que

$$\hat{Q}|\psi_0\rangle = e^{i\delta}|\psi_0\rangle \quad (17.19)$$

siempre que $|\psi_0\rangle$ sea un estado único de energía definida. Consideren el nuevo estado $|\psi'_0\rangle$ que obtienen al operar con $\hat{Q}$. Si la física es simétrica, entonces $|\psi'_0\rangle$ debe tener la misma energía que $|\psi_0\rangle$. Pero nosotros hemos considerado una situación en la que *hay solamente un* estado de esa energía, digamos $|\psi_0\rangle$, de modo que $|\psi'_0\rangle$ debe ser el mismo estado —solamente puede diferir en la fase—. Éste es el enunciado físico.

Llegamos al mismo resultado con nuestras matemáticas. Nuestra definición de simetría es la ecuación (17.10) o la ecuación (17.11) (válida para cualquier estado ψ),

$$\hat{H}\hat{Q}|\psi\rangle = \hat{Q}\hat{H}|\psi\rangle. \quad (17.20)$$

† En otros libros pueden encontrar fórmulas con signos diferentes; probablemente están usando una definición diferente para los ángulos.

Pero estamos considerando solamente a un estado $|\psi_0\rangle$, que es un estado de energía definida, de modo que $\hat{H}|\psi_0\rangle = E|\psi_0\rangle$. Como E es sólo un número que pasa al otro lado de $\hat{Q}$ si queremos, tenemos

$$\hat{Q}\hat{H}|\psi_0\rangle = \hat{Q}E|\psi_0\rangle = E\hat{Q}|\psi_0\rangle.$$

Y así

$$\hat{H}\{\hat{Q}|\psi_0\rangle\} = E\{\hat{Q}|\psi_0\rangle\}. \quad (17.21)$$

Por lo que $|\psi'_0\rangle = \hat{Q}|\psi_0\rangle$ es también un estado de energía definida de $\hat{H}$ —y con la misma E —. Pero, según nuestra hipótesis, hay sólo un estado de esa clase; por lo que debe ser $|\psi'_0\rangle = e^{i\delta}|\psi_0\rangle$.

Lo que acabamos de probar es válido para cualquier operador $\hat{Q}$ que sea un operador simétrico del sistema físico. Por lo tanto, en una situación en la que consideremos solamente fuerzas eléctricas e interacciones fuertes —y no desintegraciones β — de modo que la simetría de inversión es una aproximación permitida, tenemos que $\hat{P}|\psi\rangle = e^{i\delta}|\psi\rangle$. Pero también hemos visto que $e^{i\delta}$ debe ser $+1$ o -1 . Por lo tanto, cualquier estado de energía definida (que no sea degenerado) tiene paridad par o impar.

17-3 Las leyes de conservación

Volvemos ahora a otro ejemplo interesante de operación: una rotación. Consideramos el caso especial de un operador que rota un sistema atómico en un ángulo ϕ alrededor del eje z . Llamaremos a este operador† $\hat{R}_z(\phi)$. Supondremos que tenemos una situación física donde no hay influencias alineadas con los ejes x y y . Cualquier campo eléctrico, o campo magnético, se supondrá paralelo al eje z ,‡ así no habrá cambio en las condiciones *externas* si rotamos el sistema completo alrededor del eje z . Por ejemplo, si tenemos un átomo en el espacio vacío, y giramos el átomo alrededor del eje z un ángulo ϕ , tenemos el mismo sistema físico.

Ahora bien, hay *estados especiales* que tienen la propiedad de que tal operación produce un nuevo estado, el cual es igual al estado original multiplicado por algún factor de fase. Hagamos un rápido paréntesis para demostrarles que, cuando esto es verdad, el cambio de fase debe ser siempre proporcional al ángulo ϕ . Supongan que rotan dos veces en un ángulo ϕ . Esto es lo mismo que rotar en un ángulo 2ϕ . Si una rotación de ϕ tiene el efecto de multiplicar al estado $|\psi_0\rangle$ por una fase $e^{i\delta}$, de modo que

$$\hat{R}_z(\phi)|\psi_0\rangle = e^{i\delta}|\psi_0\rangle,$$

dos rotaciones sucesivas de ese tipo multiplicarían al estado por el factor $(e^{i\delta})^2 = e^{i2\delta}$, puesto que

$$\hat{R}_z(\phi)\hat{R}_z(\phi)|\psi_0\rangle = \hat{R}_z(\phi)e^{i\delta}|\psi_0\rangle = e^{i\delta}\hat{R}_z(\phi)|\psi_0\rangle = e^{i\delta}e^{i\delta}|\psi_0\rangle.$$

El cambio de fase δ debe ser proporcional a ϕ .¶ Estamos considerando entonces los estados especiales $|\psi_0\rangle$, para los que

$$\hat{R}_z(\phi)|\psi_0\rangle = e^{im\phi}|\psi_0\rangle, \quad (17.22)$$

donde m es algún número real.

También conocemos el hecho notable de que *si* el sistema es simétrico frente a una rotación alrededor de z *y si además* sucede que el sistema original tiene la propiedad de cumplir con (17.22), entonces también tendrá la misma propiedad

† Más precisamente, definiremos a $\hat{R}_z(\phi)$ como una rotación del sistema físico en $-\phi$ alrededor del eje z , lo cual es lo mismo que rotar el sistema de coordenadas en $+\phi$.

‡ Siempre podemos escoger z en la dirección del campo, cuando haya solamente un campo y no cambie su dirección.

¶ Para una demostración más elegante, deberíamos hacer este procedimiento para pequeñas rotaciones ϵ . Como cualquier ángulo ϕ es la suma de un número n adecuado de éstas, $\phi = n\epsilon$, $\hat{R}_z(\phi) = [\hat{R}_z(\epsilon)]^n$ y el cambio total de fase será n por el ángulo pequeño ϵ , y es, por lo tanto, proporcional a ϕ .

más tarde. Así, el número m es muy importante. Si conocemos inicialmente su valor, conocemos su valor al final del juego. Es un número que se *conserva* — m es una *constante del movimiento*—. La razón por la que sacamos a relucir al número m es porque no tiene nada que ver con algún ángulo especial ϕ y, también, porque corresponde a algo en la mecánica clásica. En la mecánica *cuántica escogemos* llamar a $m\hbar$ —para estados como $|\psi_0\rangle$ — al *momentum angular respecto al eje z* . Si hacemos esto, encontramos que en el límite de sistemas grandes, la misma cantidad es igual a la componente z del momentum angular de la mecánica clásica. Así, si tenemos un estado para el cual una rotación alrededor del eje z sólo produce un factor de fase $e^{im\phi}$, entonces tenemos un estado de momentum angular definido respecto a ese eje —y el momentum angular se conserva—. Es $m\hbar$ ahora y siempre. Naturalmente, pueden rotar alrededor de cualquier eje, y obtienen la conservación del momentum angular para los diversos ejes. Ven que la conservación del momentum angular se relaciona con el hecho de que cuando rotan un sistema, obtienen el mismo estado sólo con un nuevo factor de fase.

Nos gustaría mostrarles lo general que es esta idea. Por eso la aplicaremos a otras dos leyes de conservación, las cuales tienen correspondencia exacta, en los conceptos físicos, con la conservación del momentum angular. En física clásica tenemos también la conservación del momentum y la conservación de la energía, y es interesante ver que ambas se relacionan del mismo modo con alguna simetría física.

Supongan que tenemos un sistema físico —un átomo, algún núcleo complicado, o una molécula, o cualquier cosa— y que no importa si tomamos el sistema completo y lo movemos a un lugar diferente. De este modo, tenemos un hamiltoniano con la propiedad de que depende solamente de las *coordenadas internas*, y no depende de la *posición absoluta* en el espacio. En estas circunstancias, hay una operación de simetría especial que podemos realizar, una traslación en el espacio. Definamos a $\hat{D}_x(a)$ como la operación de desplazamiento una distancia a en la dirección x . Luego, podemos aplicar esta operación a cualquier estado y obtener un estado nuevo. Pero, de nuevo, puede haber estados muy especiales, que tienen la propiedad de que, cuando los desplazan una distancia a en la dirección del eje x , obtienen el mismo estado, excepto por un factor de fase. También es posible demostrar, como hicimos antes, que cuando esto sucede, la fase debe ser proporcional a a . Por lo que podemos escribir para estos estados especiales $|\psi_0\rangle$

$$\hat{D}_x(a) |\psi_0\rangle = e^{ika} |\psi_0\rangle. \quad (17.23)$$

El coeficiente k , cuando está multiplicado por $\hbar$, se llama *la componente x del momentum*. La razón de que se le llame así es que este número es numéricamente igual al momentum clásico p_x , cuando tenemos un sistema grande. El enunciado general es: si el hamiltoniano no varía cuando se desplaza el sistema, y si el estado inicia con un momentum definido en la dirección x , entonces el momentum en la dirección x permanecerá constante al transcurrir el tiempo. El momentum total de un sistema antes y después de una colisión —o después de explosiones, o de lo que sea— será el mismo.

Hay otra operación que es muy parecida a un desplazamiento en el espacio: un retraso en el tiempo. Supongan que tenemos una situación física donde no hay *nada externo* que dependa del tiempo, y hacemos que algo comience en un estado dado en un cierto momento y lo dejamos marchar. Ahora bien, si dejásemos que la misma cosa (en otro experimento) comenzase dos segundos más tarde —o digamos, retasado un tiempo τ — y si nada en las condiciones externas depende del tiempo absoluto, la evolución sería la misma, y el estado final sería el mismo que el otro estado final, sólo que llegará un tiempo τ más tarde. En esas circunstancias, también podemos encontrar estados especiales, que tienen la propiedad de que la evolución en el tiempo tiene las características especiales de que el estado retardado es precisamente el anterior, multiplicado por un factor de fase. Una vez más, es claro que para estos estados especiales, el cambio de fase debe ser proporcional a τ . Podemos escribir

$$\hat{D}_t(\tau) |\psi_0\rangle = e^{-i\omega\tau} |\psi_0\rangle. \quad (17.24)$$

Por convención, se usa el signo negativo al definir ω ; con esta convención, $\omega\hbar$ es la *energía* del sistema, y *ésta se conserva*. Así, un estado de energía definida es tal que, cuando se lo desplaza τ en el tiempo, se reproduce a sí mismo multiplicado por $e^{-i\omega\tau}$. (Esto es lo que dijimos antes, cuando definimos un estado cuántico de energía definida, por lo que somos consistentes con nosotros mismos.) Esto significa que, si un sistema está en un estado de energía definida, y que si el hamiltoniano no depende de t , pase lo que pase, el sistema tendrá la misma energía en tiempos posteriores.

Observemos la relación entre las leyes de conservación y la simetría del mundo. Simetría respecto a desplazamientos en el tiempo implica conservación de energía; simetría respecto a la posición en x , y o z implica conservación de dicha componente del momentum. Simetría respecto a rotaciones alrededor de los ejes x , y y z implica conservación de las componentes x , y y z del momentum angular. Simetría respecto a reflexión implica conservación de la paridad. Simetría respecto al intercambio de dos electrones implica conservación de algo a lo que no hemos dado nombre, y así sucesivamente. Algunos de estos principios tienen sus análogos clásicos y otros no. Hay más leyes de conservación en la mecánica cuántica, de las que son útiles en la mecánica clásica —o, por lo menos, las que se usan comúnmente—.

Con objeto de que puedan leer otros libros de mecánica cuántica, debemos hacer un pequeño paréntesis —para describir la notación que la gente usa—. La operación de desplazamiento respecto al tiempo es, por supuesto, la operación $\hat{U}$ de que hablamos antes:

$$\hat{D}_t(\tau) = \hat{U}(t + \tau, t). \quad (17.25)$$

A muchos les gusta discutir todo en términos de desplazamientos *infinitesimales* en el tiempo, o en términos de desplazamientos infinitesimales en el espacio, o en términos de rotaciones de ángulos infinitesimales. Como cualquier desplazamiento o ángulo distinto de cero se puede formar por acumulación de una sucesión de desplazamientos o ángulos infinitesimales, con frecuencia es más fácil analizar primero el caso infinitesimal. El operador de un desplazamiento infinitesimal Δt en el tiempo es —como lo definimos en el capítulo 8—

$$\hat{D}_t(\Delta t) = 1 - \frac{i}{\hbar} \Delta t \hat{H}. \quad (17.26)$$

Entonces $\hat{H}$ es análogo a la cantidad clásica que llamamos energía, porque si $\hat{H}|\psi\rangle$ resulta ser una constante por $|\psi\rangle$, digamos $\hat{H}|\psi\rangle = E|\psi\rangle$, entonces esta constante es la energía del sistema.

Lo mismo se hace para las otras operaciones. Si tenemos un pequeño desplazamiento en x , digamos Δx , un estado $|\psi\rangle$, *en general*, pasará a algún otro estado $|\psi'\rangle$. Podemos escribir

$$|\psi'\rangle = \hat{D}_x(\Delta x)|\psi\rangle = \left(1 + \frac{i}{\hbar} \hat{p}_x \Delta x\right)|\psi\rangle, \quad (17.27)$$

puesto que cuando Δx tiende a cero, $|\psi'\rangle$ se debe convertir simplemente en $|\psi\rangle$ o sea $\hat{D}_x(0) = 1$, y para Δx pequeños, la variación de $\hat{D}_x(\Delta x)$ de 1 debe ser proporcional a Δx . Definido así, el operador $\hat{p}_x$ se llama operador momentum —naturalmente, para la componente x —.

Por razones idénticas, la gente suele escribir para rotaciones pequeñas

$$\hat{R}_z(\Delta\phi)|\psi\rangle = \left(1 + \frac{i}{\hbar} \hat{J}_z \Delta\phi\right)|\psi\rangle \quad (17.28)$$

y llaman $\hat{J}_z$ al operador de la componente z del momentum angular. Para los estados especiales para los que $\hat{R}_z(\phi)|\psi_0\rangle = e^{im\phi}|\psi_0\rangle$, para cualquier ángulo pequeño —digamos $\Delta\phi$ — podemos expandir el lado derecho hasta el primer orden en $\Delta\phi$, y obtenemos

$$\hat{R}_z(\Delta\phi)|\psi_0\rangle = e^{im\Delta\phi}|\psi_0\rangle = (1 + im\Delta\phi)|\psi_0\rangle.$$

Comparando esto con la definición de $\hat{J}_z$ en la ecuación (17.28), obtenemos que

$$\hat{J}_z |\psi_0\rangle = m\hbar |\psi_0\rangle. \quad (17.29)$$

En otras palabras, si aplican $\hat{J}_z$ a un estado *con momentum angular definido* respecto al eje z , obtienen $m\hbar$ por el mismo estado, donde $m\hbar$ es la cantidad de la componente z del momentum angular. Es bastante similar a aplicar $\hat{H}$ a un estado de energía definida para obtener $E|\psi\rangle$.

Ahora nos gustaría hacer algunas aplicaciones de las ideas de la conservación del momentum angular —para mostrarles cómo funcionan—. La cuestión es que son realmente muy simples. Ya sabían que el momentum angular se conserva. Lo único que tienen que recordar de este capítulo es que si un estado $|\psi_0\rangle$ tiene la propiedad de que, sometido a una rotación en un ángulo ϕ alrededor del eje z , se convierte en $e^{im\phi}|\psi_0\rangle$; tiene una componente z del momentum angular igual a $m\hbar$. Esto es todo lo que necesitamos para hacer algunas cosas interesantes.

17-4 Luz polarizada

Primero nos gustaría comprobar una idea. En la sección 11-4 vimos que cuando luz polarizada RHC se observaba desde un sistema girado un ángulo ϕ alrededor del eje $z^\dagger$, ésta era multiplicada por $e^{i\phi}$. ¿Esto significa que los fotones de luz polarizados circularmente a la derecha llevan un momentum angular de una unidad ‡ en la dirección del eje z ? *Así es en verdad.* Y también significa que si tenemos un haz de luz que contenga un gran número de fotones, todos polarizados circularmente del mismo modo —como tendríamos en un haz clásico— éste llevará un momentum angular. Si la energía total que lleva el haz en un cierto tiempo es W , entonces hay $N = W/\hbar\omega$ fotones. Cada uno lleva el momentum angular $\hbar$, por lo que el momentum angular total es

$$J_z = N\hbar = \frac{W}{\omega}. \quad (17.30)$$

¿Podemos probar clásicamente que la luz polarizada circularmente a la derecha lleva una energía y un momentum angular proporcional a W/ω ? Esto sería una proposición clásica, si todo fuese cierto. Aquí tenemos un caso donde podemos ir de la situación cuántica a la clásica. Debemos ver si la física clásica concuerda. Nos dará una idea de si tenemos derecho a llamar momentum angular a m . Recuerden lo que es, clásicamente, luz polarizada circularmente a la derecha. Se describe mediante un campo eléctrico con una componente x oscilatoria y una componente y oscilatoria desfasadas 90° , de modo que el vector eléctrico resultante $\mathcal{E}$ se mueve en un círculo —como se muestra en la figura 17-5(a)—. Ahora supongan que esa luz incide sobre una pared que va a absorberla —por lo menos en parte— y consideren a un átomo en la pared de acuerdo con la física clásica. Hemos descrito con frecuencia el movimiento del electrón en el átomo como un oscilador armónico, el cual puede ser inducido a oscilar mediante un campo eléctrico externo. Supondremos que el átomo es isótropo, de modo que pueda oscilar igualmente bien en las direcciones x o y . Así, en luz polarizada circularmente, los desplazamientos en x y en y son iguales, pero uno va 90° detrás del otro. El resultado neto es que el electrón se mueve en un círculo, como muestra la figura 17-5(b). El electrón sufre un desplazamiento $\mathbf{r}$ de su posición de equilibrio en el origen, y rota con algún retraso de fase respecto al vector $\mathcal{E}$. La relación entre $\mathbf{r}$ y $\mathcal{E}$ bien podría ser como la que muestra la figura 17-5(b). Al pasar el tiempo, el campo eléctrico rota y el desplazamiento rota con la misma frecuencia, por lo que su orientación relativa permanece constante. Ahora veamos el trabajo que se ha efectuado sobre el electrón. La rapidez con que

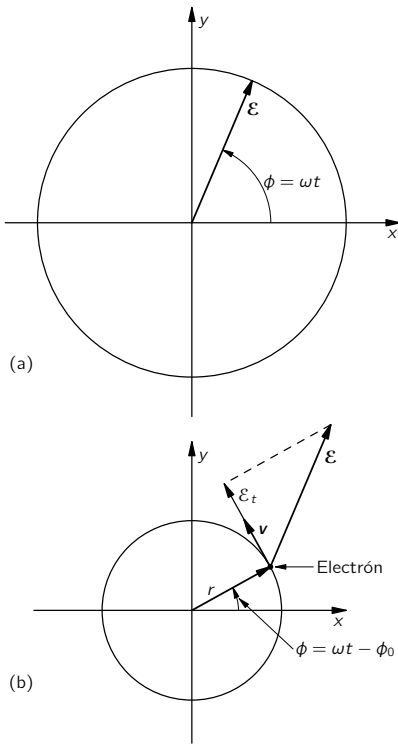


Figura 17-5. (a) El campo eléctrico $\mathcal{E}$ en una onda de luz circularmente polarizada. (b) El movimiento de un electrón dirigido por la luz circularmente polarizada.

† ¡Lo siento! Este ángulo es el negativo del que usamos en la sección 11-4.

‡ En general, es muy conveniente medir el momentum angular de sistemas atómicos en unidades $\hbar$. Así, pueden decir que una partícula con espín un medio tiene un momentum angular igual a $\pm 1/2$ respecto a cualquier eje. O, en general, que la componente z del momentum angular es m . No hay necesidad de repetir la $\hbar$ todo el tiempo.

la energía pasa al electrón es v , su velocidad, por la componente de $q\mathcal{E}$ paralela a la velocidad:

$$\frac{dW}{dt} = q\mathcal{E}_t v. \quad (17.31)$$

Pero observen, hay un momentum angular que se ha vertido en el electrón, ya que siempre hay un torque respecto al origen. El torque es $q\mathcal{E}_t r$, que debe ser igual a la rapidez de variación del momentum angular dJ_z/dt :

$$\frac{dJ_z}{dt} = q\mathcal{E}_t r. \quad (17.32)$$

Recordando que $v = \omega r$, tenemos que

$$\frac{dJ_z}{dW} = \frac{1}{\omega}.$$

Por lo tanto, si integramos el momentum angular total que se absorbe, resulta ser proporcional a la energía total —la constante de proporcionalidad es $1/\omega$, lo que está de acuerdo con la ecuación (17.30)—. La luz sí lleva un momentum angular —1 unidad (por $\hbar$) si está polarizada circularmente a la derecha en la dirección del eje z , y -1 unidad en la dirección del eje z si está polarizada circularmente a la izquierda—.

Ahora hagamos la siguiente pregunta: si la luz está linealmente polarizada en la dirección x , ¿cuál es su momentum angular? La luz polarizada en la dirección x se puede representar como una superposición de luz RHC y luz LHC. Por lo tanto, hay una cierta amplitud para que el momentum angular sea $+\hbar$, y otra amplitud para que el momentum angular sea $-\hbar$; por lo tanto, no tiene un momentum angular *definido*. Tiene una amplitud para aparecer con $+\hbar$, y una amplitud igual para aparecer con $-\hbar$. La interferencia de estas dos amplitudes produce la polarización lineal, pero tiene probabilidades *iguales* para aparecer con más o menos una unidad de momentum angular. Las medidas macroscópicas hechas en un haz de luz polarizada linealmente mostrarán que lleva momentum angular cero, debido a que en un gran número de fotones hay números casi iguales de fotones RHC y LHC contribuyendo con cantidades opuestas de momentum angular —el momentum angular promedio es cero—. En la teoría clásica no encuentran momentum angular, a no ser que haya alguna polarización circular.

Hemos dicho que cualquier partícula con espín uno puede tener tres valores para J_z , a saber $+1$, 0 , -1 (los tres estados que encontramos en el experimento de Stern-Gerlach). Pero la luz es excéntrica, tiene solamente dos estados. No tiene el caso cero. Esta extraña carencia se debe a que la luz no puede permanecer quieta. Para una partícula de espín j que está quieta, debe haber $2j + 1$ estados posibles con valores de j_z que varían de 1 en 1 , desde $-j$ hasta $+j$. Pero resulta que para algo con espín j y masa cero, solamente existen los estados con componentes $+j$ y $-j$ en la dirección del movimiento. Por ejemplo, la luz no tiene tres estados, sino solamente dos —aunque un fotón sea un objeto de espín uno—. ¿Cómo es esto consistente con nuestras demostraciones anteriores —basadas en lo que sucede frente a rotaciones en el espacio— de que son necesarios tres estados para partículas de espín uno? Para una partícula en reposo, se pueden hacer rotaciones alrededor de cualquier eje sin que cambie el estado de momentum. Las partículas con masa en reposo cero (como los fotones y neutrinos) no pueden estar en reposo; solamente las rotaciones alrededor de la dirección del movimiento no cambian el estado de momentum. Los argumentos sobre rotaciones alrededor de sólo un eje son insuficientes para probar que se necesita tres estados, dado que uno de ellos varía como $e^{i\phi}$ frente a rotaciones en un ángulo ϕ .†

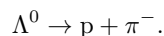
Una observación más. Para una partícula de masa en reposo cero, en general, *sólo uno* de los dos estados de espín respecto a la línea de movimiento ($+j$, $-j$)

† Hemos tratado de encontrar, por lo menos, una demostración de que la componente del momentum angular en la dirección de movimiento debe ser, para una partícula de masa cero, un múltiplo entero de $\hbar/2$ —y no algo como $\hbar/3$ —. Aun usando todas las propiedades de transformación de Lorentz y qué sé yo, fracasamos. Puede ser que no sea verdad. Tendremos que hablar sobre esto con el Profesor Wigner, quien sabe todo sobre estas cosas.

es realmente necesario. Para neutrinos —que son partículas con espín un medio— solamente los estados con componente de momentum angular *opuesto* a la dirección de movimiento ($-\hbar/2$) existen en la naturaleza [y solamente *en la dirección* del movimiento ($+\hbar/2$) para antineutrinos]. Cuando un sistema tiene simetría de inversión (por lo tanto, se conserva la paridad, como ocurre con la luz) se requieren ambas componentes ($+j$ y $-j$).

17-5 La desintegración de la Λ^0

Queremos dar ahora un ejemplo de cómo usar el teorema de conservación del momentum angular en un problema cuántico específico. Consideramos la rotura de la partícula lambda (Λ^0), que se desintegra en un protón y un mesón π^- mediante una interacción “débil”:



Supongan que sabemos que el pion tiene espín cero, que el protón tiene espín un medio y que la Λ^0 tiene espín un medio. Nos gustaría resolver el siguiente problema: consideren que se va a producir una Λ^0 de modo que va a estar polarizado completamente —queremos dar a entender que su espín está, por ejemplo, “hacia arriba”, respecto a algún eje z escogido convenientemente— [véase la figura 17-6(a)]. La pregunta es, ¿con qué probabilidad se desintegrará de modo que el protón salga formando un ángulo θ respecto al eje z —como en la figura 17-6(b)—? En otras palabras, ¿cuál es la distribución angular de las desintegraciones? Consideraremos esta desintegración en el sistema de coordenadas en el cual Λ^0 está en reposo —mediremos los ángulos en este sistema en reposo; luego, si lo deseamos, lo podremos transformar a otro sistema—.

Comencemos considerando el caso especial en que el protón es emitido en un ángulo sólido pequeño $\Delta\Omega$ en dirección del eje z (figura 17-7). Antes de la desintegración, tenemos una Λ^0 con espín “hacia arriba”, como en la parte (a) de la figura. Después de un corto tiempo —por razones desconocidas hasta hoy, pero que se relacionan con las desintegraciones débiles— la Λ^0 explota en un protón y un pion. Supongan que el protón va hacia arriba en dirección del eje $+z$. Entonces, por la conservación del momentum, el pion debe ir hacia abajo. Como el protón es una partícula con espín un medio, su espín debe estar “hacia arriba” o “hacia abajo” —son, en principio, las dos posibilidades (b) y (c) que muestra la figura—. Sin embargo, la conservación del momentum angular requiere que el protón tenga espín “hacia arriba”. Esto se puede ver mejor a través del siguiente argumento. Una partícula que se mueve en la dirección del eje z no contribuye con ningún momentum angular respecto a este eje en virtud de su movimiento; por lo tanto, solamente los espines pueden contribuir a J_z . El momentum angular del espín respecto al eje z es $+\hbar/2$ antes de la desintegración, por lo que debe ser $+\hbar/2$ después de la misma. Podemos decir que, puesto que el pion no tiene espín, el espín del protón debe ser “hacia arriba”.

Si les preocupa que esta clase de argumentos tal vez no sean válidos en la mecánica cuántica, podemos tomarnos un momento para demostrarles que lo son. El estado inicial (antes de la desintegración), al que llamaremos $|\Lambda^0, \text{espín } +z\rangle$, tiene la propiedad de que, si se rota alrededor del eje z en un ángulo ϕ , el vector de estado queda multiplicado por el factor de fase $e^{i\phi/2}$. (En el sistema rotado, el vector de estado es $e^{i\phi/2}|\Lambda^0, \text{espín } +z\rangle$.) Esto es lo que queremos dar a entender por espín “hacia arriba” para una partícula con espín un medio. Como el comportamiento de la naturaleza no depende de nuestra elección de ejes, el estado final (protón más pion) debe tener la misma propiedad. Podríamos escribir el estado final como, por ejemplo,

$$|\text{protón en dirección } +z, \text{ espín } +z; \text{ pion en dirección } -z\rangle.$$

Pero realmente no tenemos por qué especificar el movimiento del pion, ya que en el sistema que hemos elegido, el pion siempre se mueve en dirección opuesta al protón; podemos simplificar nuestra descripción del estado final a

$$|\text{protón en dirección } +z, \text{ espín } +z\rangle.$$

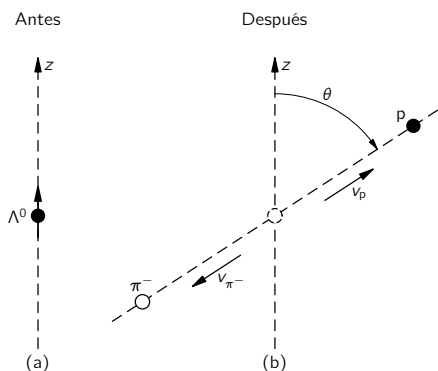


Figura 17-6. Una Λ^0 con espín “hacia arriba” se desintegra en un protón y un pion (en el sistema CM). ¿Cuál es la probabilidad para que el protón salga con un ángulo θ ?

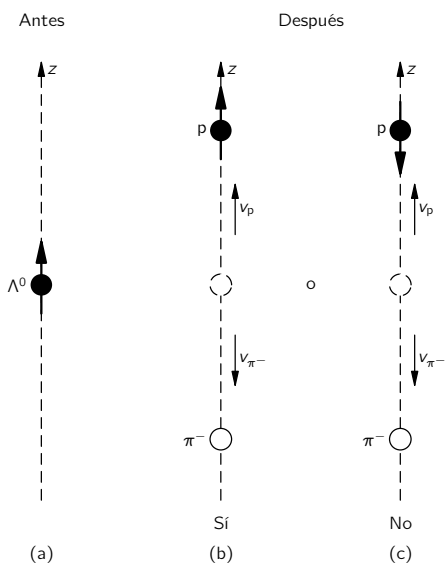


Figura 17-7. Dos posibilidades para la desintegración de un Λ^0 con espín “hacia arriba” con el protón yendo en la dirección del eje $+z$. Solamente (b) conserva el momentum angular.

Ahora bien, ¿qué sucede a este vector de estado si rotamos las coordenadas alrededor del eje z en un ángulo ϕ ?

Como el protón y el pión se mueven en la dirección del eje z , entonces su movimiento no se ve afectado por la rotación. (Es por esto que escogimos este caso especial; no podíamos hacer el argumento de otro modo.) También, nada sucede al pión porque tiene espín cero. El protón, sin embargo, tiene espín un medio. Si su espín está “hacia arriba”, entonces éste contribuirá con un cambio de fase de $e^{i\phi/2}$ en respuesta a la rotación. (Si su espín estuviese “hacia abajo”, el cambio de fase debido al protón sería $e^{-i\phi/2}$.) Pero el cambio de fase con la rotación antes y después de la excitación debe ser el mismo si se tiene que conservar el momentum angular. (Y así será, ya que no hay influencias externas en el hamiltoniano.) Así, la única posibilidad es que el espín del protón esté “hacia arriba”. Si el protón va hacia arriba, su espín también debe ser “hacia arriba”.

Concluimos, entonces, que la conservación del momentum angular permite el proceso de la parte (b) de la figura 17-7, pero no permite el proceso de la parte (c). Como sabemos que la desintegración tiene lugar, hay una amplitud para el proceso (b) —el protón que va hacia arriba con el espín “hacia arriba”—. Sea a la amplitud para que la desintegración ocurra de ese modo en cualquier intervalo infinitesimal de tiempo.†

Veamos ahora lo que sucedería si el espín de la Λ^0 estuviese inicialmente “hacia abajo”. De nuevo, nos preguntamos por las desintegraciones en las que el protón va hacia arriba en dirección del eje z , como lo muestra la figura 17-8. Aprenderán que, en este caso, el protón debe tener espín “hacia abajo” si se conserva el momentum angular. Sea b la amplitud para tal desintegración.

No podemos decir nada más acerca de las amplitudes a y b . Dependen de la maquinaria interna de la Λ^0 y de las desintegraciones débiles, y nadie sabe aún cómo calcularlas. Tendremos que obtenerlas experimentalmente. Pero justo con estas dos amplitudes *podemos* deducir todo lo que queramos conocer acerca de la distribución angular de la desintegración. Lo único es que tenemos que ser muy cuidadosos de definir completamente los estados que estemos tratando.

Queremos conocer la probabilidad para que el protón salga formando un ángulo θ respecto al eje z (dentro de un pequeño ángulo sólido $\Delta\Omega$) como muestra la figura 17-6. Pongamos un nuevo eje z en esta dirección y llamémosle z' . Sabemos cómo analizar lo que sucede en este eje. Respecto a este nuevo eje, la Λ^0 ya no tiene su espín “hacia arriba”, sino tiene una cierta amplitud para tener su espín “hacia arriba” y otra para tener su espín “hacia abajo”. Ya las calculamos en el capítulo 6 y lo hicimos nuevamente en el capítulo 10, ecuación (10.30). La amplitud para tener espín “hacia arriba” es $\cos\theta/2$, y la amplitud para tener espín “hacia abajo” es $\sin\theta/2$. Cuando el espín de la Λ^0 está “hacia arriba” en la dirección de z' , emitirá un protón en la dirección $+z'$ con la amplitud a . Así, la amplitud para encontrar un protón con espín “hacia arriba” en la dirección z' es

$$a \cos \frac{\theta}{2}. \quad (17.33)$$

Del mismo modo, la amplitud para encontrar un protón con espín “hacia abajo” en la dirección positiva de z' es

$$-b \sin \frac{\theta}{2}. \quad (17.34)$$

La figura 17-9 muestra los dos procesos a los que se refieren estas amplitudes.

Ahora hagamos la siguiente pregunta sencilla. Si la Λ^0 tiene espín “hacia arriba” en dirección del eje z , ¿cuál es la probabilidad para que el protón de la

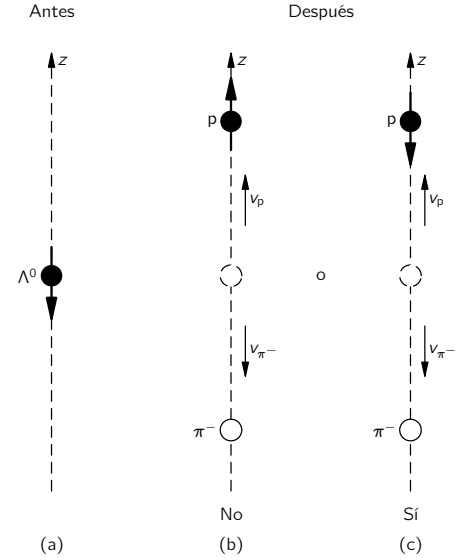


Figura 17-8. La desintegración en dirección del eje z para una partícula Λ^0 con espín “hacia abajo”.

† Ahora estamos suponiendo que la maquinaria de la mecánica cuántica les es suficientemente familiar para que podamos hablar de cosas de un modo físico sin tomarnos la molestia de escribir todos sus detalles matemáticos. En caso de que no les resulte claro lo que estamos diciendo, hemos puesto algunos de los detalles omitidos en una nota al final de la sección.

‡ Hemos escogido z' en el plano xz y usamos los elementos matriciales para $R_y(\theta)$. Ustedes obtendrían la misma respuesta para cualquier otra elección.

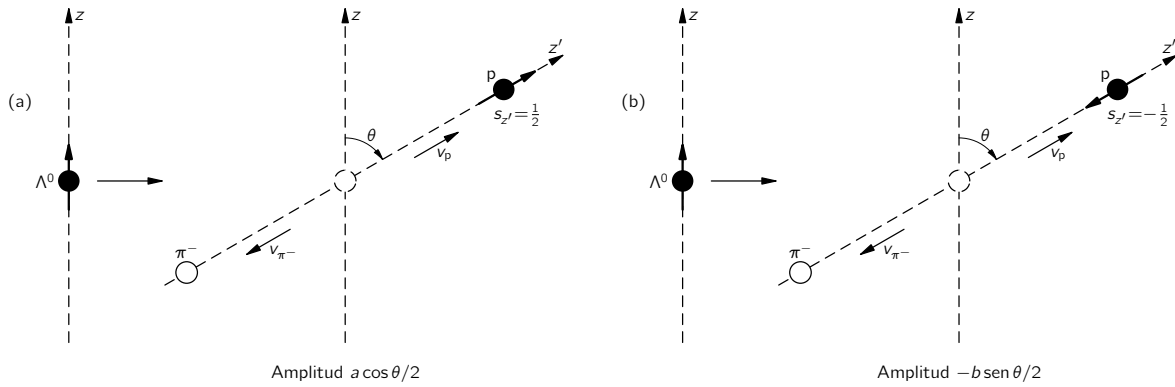


Figura 17-9. Dos estados posibles de desintegración para la Λ^0 .

desintegración salga formando el ángulo θ ? Los dos estados de espín (“hacia arriba” o “hacia abajo” en dirección de z') son distinguibles, a pesar de que no los vamos a considerar. Por lo cual, para obtener la probabilidad, elevamos las amplitudes al cuadrado y sumamos. La probabilidad $f(\theta)$ para encontrar un protón en un pequeño ángulo sólido $\Delta\Omega$ en θ es

$$f(\theta) = |a|^2 \cos^2 \frac{\theta}{2} + |b|^2 \sin^2 \frac{\theta}{2}. \quad (17.35)$$

Recordando que $\sin^2 \theta/2 = \frac{1}{2}(1 - \cos \theta)$ y que $\cos^2 \theta/2 = \frac{1}{2}(1 + \cos \theta)$, podemos escribir $f(\theta)$ en la forma

$$f(\theta) = \left(\frac{|a|^2 + |b|^2}{2} \right) + \left(\frac{|a|^2 - |b|^2}{2} \right) \cos \theta. \quad (17.36)$$

La distribución angular tiene la forma

$$f(\theta) = \beta(1 + \alpha \cos \theta). \quad (17.37)$$

La probabilidad tiene una parte que es independiente de θ y otra que varía linealmente con $\cos \theta$. Midiendo la distribución angular podemos obtener α y β y, por lo tanto, $|a|$ y $|b|$.

Ahora podemos responder muchas preguntas. ¿Estamos solamente interesados en protones con espín “hacia arriba” en la dirección del eje z antiguo? Cada uno de los términos de (17.33) y de (17.34) dará una amplitud para encontrar un protón con espín “hacia arriba” y con espín “hacia abajo” respecto al eje z' ($+z'$ y $-z'$). El espín “hacia arriba” respecto al antiguo eje $|+z\rangle$, se puede expresar en términos de los estados base $|+z'\rangle$ y $|-z'\rangle$. Entonces podemos combinar las dos amplitudes (17.33) y (17.34) con los coeficientes apropiados ($\cos \theta/2$ y $-\sin \theta/2$) para obtener la amplitud total

$$\left(a \cos^2 \frac{\theta}{2} + b \sin^2 \frac{\theta}{2} \right).$$

Su cuadrado es la probabilidad para que el protón salga formando el ángulo θ con espín igual al de la Λ^0 (“hacia arriba” en la dirección del eje z).

Si se conservase la paridad, podríamos decir algo más. La desintegración de la figura 17-8 es justamente la reflexión —en, digamos, el plano yz de la desintegración— de la figura 17-7.† Si la paridad se conservase, b tendría que ser igual a a o a $-a$. Entonces el coeficiente α de (17.37) sería cero, y la desintegración tendría la misma probabilidad para ocurrir en todas direcciones.

No obstante, los resultados experimentales muestran que *hay* una asimetría en la desintegración. La distribución angular medida es como $\cos \theta$ según predijimos

† Recordando que el espín es un vector axial y no se invierte en la reflexión.

—y no como $\cos^2 \theta$ o cualquier otra potencia—. De hecho, ya que la distribución angular tiene esta forma, podemos deducir de estas mediciones que el espín de la Λ^0 es $1/2$. Además, vemos que la paridad no se conserva. En realidad, se ha encontrado experimentalmente que el coeficiente α es -0.62 ± 0.05 , de modo que b es casi dos veces mayor que a . La falta de simetría frente a una reflexión es evidente.

Se dan cuenta de cuántas cosas podemos obtener a partir de la conservación del momentum angular. En el próximo capítulo daremos algunos otros ejemplos.

Nota entre paréntesis. Por la amplitud a en esta sección, entendemos la amplitud para que el estado |protón en la dirección $+z$, espín $+z$ ⟩ sea generado en un tiempo infinitesimal dt a partir del estado $|\Lambda, \text{espín } +z\rangle$, o, en otras palabras, que

$$\langle \text{protón en dirección } +z, \text{espín } +z | H | \Lambda, \text{espín } +z \rangle = i\hbar a, \quad (17.38)$$

donde H es el hamiltoniano del mundo —o por lo menos, de aquello que es responsable por la desintegración de la Λ —. La conservación del momentum angular significa que el hamiltoniano debe tener la propiedad de

$$\langle \text{protón en la dirección } +z, \text{espín } -z | H | \Lambda, \text{espín } +z \rangle = 0. \quad (17.39)$$

Por la amplitud b entendemos que

$$\langle \text{protón en la dirección } +z, \text{espín } -z | H | \Lambda, \text{espín } -z \rangle = i\hbar b. \quad (17.40)$$

La conservación del momentum angular implica que

$$\langle \text{protón en la dirección } +z, \text{espín } +z | H | \Lambda, \text{espín } -z \rangle = 0. \quad (17.41)$$

Si las amplitudes escritas en (17.33) y (17.34) no resultan claras, podemos expresarlas en una forma más matemática como sigue. Por (17.33) entendemos la amplitud para que la Λ , con espín en la dirección $+z$, se desintegre en un protón que se mueve en la dirección $+z'$, con su espín también en la dirección $+z'$, es decir, la amplitud

$$\langle \text{protón en la dirección } +z', \text{espín } +z' | H | \Lambda, \text{espín } +z \rangle. \quad (17.42)$$

Por los teoremas generales de la mecánica cuántica, esta amplitud se puede escribir en la forma

$$\sum_i \langle \text{protón en la dirección } +z', \text{espín } +z' | H | \Lambda, i \rangle \langle \Lambda, i | \Lambda, \text{espín } +z \rangle, \quad (17.43)$$

donde hay que extender la suma a los estados base $|\Lambda, i\rangle$ de la partícula Λ en reposo. Como la partícula Λ tiene espín un medio, hay dos estados base que pueden estar en cualquier base de referencia que deseemos. Si usamos como estados base el espín “hacia arriba” y el espín “hacia abajo” *respecto a z'* ($+z', -z'$), la amplitud (17.43) es igual a la suma

$$\begin{aligned} & \langle \text{protón en la dirección } +z', \text{espín } +z' | H | \Lambda, +z' \rangle \langle \Lambda, +z' | \Lambda, +z \rangle \\ & + \langle \text{protón en la dirección } +z', \text{espín } +z' | H | \Lambda, -z' \rangle \langle \Lambda, -z' | \Lambda, +z \rangle. \end{aligned} \quad (17.44)$$

El primer factor del primer término es a y el primer factor del segundo término es cero —por la definición (17.38) y la ecuación (17.41), que a su vez proceden de la conservación del momentum angular—. El factor restante del primer término $\langle \Lambda, +z' | \Lambda, +z \rangle$ es justamente la amplitud para que una partícula con espín un medio, la cual tiene espín “hacia arriba” en la dirección de un eje, tenga también espín “hacia arriba” en la dirección de un eje girado el ángulo θ , que es $\cos \theta/2$ —véase la tabla 6-2—. Así, (17.44) es justamente $a \cos \theta/2$, como escribimos en (17.33). La amplitud de (17.34) se obtiene con los mismos argumentos para una partícula Λ con espín “hacia abajo”.

17-6 Resumen de las matrices de rotación

Queremos reunir en un solo lugar los diversos aspectos que hemos aprendido acerca de rotaciones para partículas con espín un medio y con espín uno —nos será cómodo para referencias futuras—. En la página 17-17 encontrarán tablas de las dos matrices de rotación $R_z(\phi)$ y $R_y(\theta)$ para partículas con espín un medio, con espín uno y para fotones (partículas con espín uno y con masa en reposo cero). Para cada espín daremos los términos de la matriz $\langle j | R | i \rangle$ para rotaciones alrededor del eje z o del eje y . Son, naturalmente, exactamente equivalentes a las amplitudes como $\langle +T | 0 S \rangle$ que usamos en los capítulos anteriores. Entendemos que $R_z(\phi)$ es el estado proyectado en un nuevo sistema de coordenadas girado en el ángulo ϕ alrededor del eje z —usando siempre la regla de la mano derecha para definir el sentido positivo de la rotación—. Por $R_y(\theta)$ decimos que los ejes de referencia están rotados en el ángulo θ alrededor del eje y . Conociendo estas dos rotaciones, pueden, naturalmente, calcular cualquier rotación arbitraria. Como de costumbre, escribimos los elementos matriciales de modo que el estado a la izquierda es un estado base del *nuevo* sistema (girado) y el estado a la derecha es un estado base del sistema viejo (sin girar). Pueden interpretar las entradas en las tablas de muchas maneras. Por ejemplo, la entrada $e^{-i\phi/2}$ en la tabla 17-1 significa que el elemento matricial $\langle - | R | - \rangle = e^{-i\phi/2}$. También significa que $\hat{R} | - \rangle = e^{-i\phi/2} | - \rangle$, o que $\langle - | \hat{R} = \langle - | e^{-i\phi/2}$. Todo es la misma cosa.

Tabla 17-1

Matrices de rotación para espín un medio

Dos estados: $|+\rangle$, “hacia arriba” en dirección del eje z , $m = +1/2$
 $|-\rangle$, “hacia abajo” en dirección del eje z , $m = -1/2$

$R_z(\phi)$	$ +\rangle$	$ -\rangle$
$\langle+ $	$e^{+i\phi/2}$	0
$\langle- $	0	$e^{-i\phi/2}$

$R_y(\theta)$	$ +\rangle$	$ -\rangle$
$\langle+ $	$\cos \theta/2$	$\sin \theta/2$
$\langle- $	$-\sin \theta/2$	$\cos \theta/2$

Tabla 17-2

Matrices de rotación para espín uno

Tres estados: $|+\rangle$, $m = +1$
 $|0\rangle$, $m = 0$
 $|-\rangle$, $m = -1$

$R_z(\phi)$	$ +\rangle$	$ 0\rangle$	$ -\rangle$
$\langle+ $	$e^{+i\phi}$	0	0
$\langle 0 $	0	1	0
$\langle- $	0	0	$e^{-i\phi}$

$R_y(\theta)$	$ +\rangle$	$ 0\rangle$	$ -\rangle$
$\langle+ $	$\frac{1}{2}(1 + \cos \theta)$	$+\frac{1}{\sqrt{2}} \sin \theta$	$\frac{1}{2}(1 - \cos \theta)$
$\langle 0 $	$-\frac{1}{\sqrt{2}} \sin \theta$	$\cos \theta$	$+\frac{1}{\sqrt{2}} \sin \theta$
$\langle- $	$\frac{1}{2}(1 - \cos \theta)$	$-\frac{1}{\sqrt{2}} \sin \theta$	$\frac{1}{2}(1 + \cos \theta)$

Tabla 17-3

Fotones

Dos estados: $|R\rangle = \frac{1}{\sqrt{2}}(|x\rangle + i|y\rangle)$, $m = +1$ (polarizado RHC)

$|L\rangle = \frac{1}{\sqrt{2}}(|x\rangle - i|y\rangle)$, $m = -1$ (polarizado LHC)

$R_z(\phi)$	$ R\rangle$	$ L\rangle$
$\langle R $	$e^{+i\phi}$	0
$\langle L $	0	$e^{-i\phi}$

Momentum angular

18-1 Radiación dipolar eléctrica

En el capítulo anterior desarrollamos el concepto de conservación del momentum angular en mecánica cuántica y mostramos cómo se podría usar para predecir la distribución angular del protón proveniente de la desintegración de la partícula Λ . Ahora queremos agregar una cantidad de ejemplos similares de las consecuencias de la conservación del momentum en sistemas atómicos. Nuestro primer ejemplo es la radiación de luz por un átomo. La conservación del momentum angular (entre otras cosas) determinará la polarización y la distribución angular de los fotones emitidos.

Supongan que tenemos un átomo en un estado excitado de momentum angular definido —digamos, con espín uno— y que hace una transición a un estado de momentum angular cero a una energía más baja, emitiendo un fotón. El problema es calcular la distribución angular y la polarización de los fotones. (Esta cuestión es casi igual al de la desintegración de la Λ^0 , excepto que tenemos partículas de espín uno en vez de partículas de espín un medio.) Como el estado superior del átomo tiene espín uno, hay tres posibilidades para la componente z de su momentum angular. El valor de m podría ser $+1$ o 0 o -1 . En nuestro ejemplo tomaremos $m = +1$. Una vez que vean cómo se hace, pueden calcular los otros casos. Suponemos que el átomo está con su momentum angular en dirección del eje $+z$ —como en la figura 18-1(a)— y preguntamos con qué amplitud emitirá luz circularmente polarizada a la derecha y hacia arriba en dirección del eje z , de modo que el átomo termine con momentum angular cero —como muestra la parte (b) de la figura—. Bueno, no conocemos la respuesta. Pero lo que sí sabemos es que la luz circularmente polarizada a la derecha tiene una unidad de momentum angular en su dirección de propagación. Así, después de que se emite el fotón, la situación tendría que ser como muestra la figura 18-1(b) —el

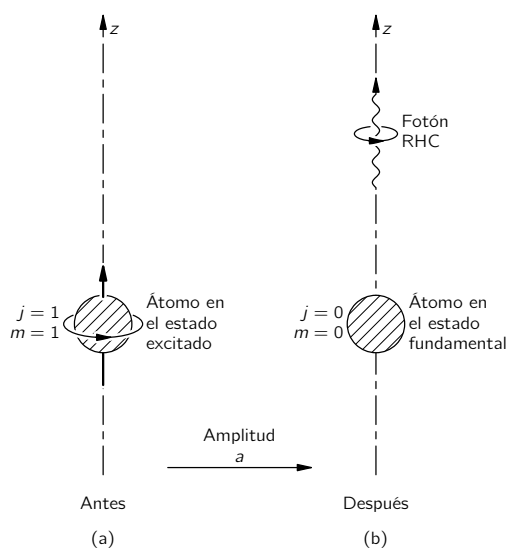


Figura 18-1. Un átomo con $m = +1$ emite un fotón RHC en dirección del eje $+z$.

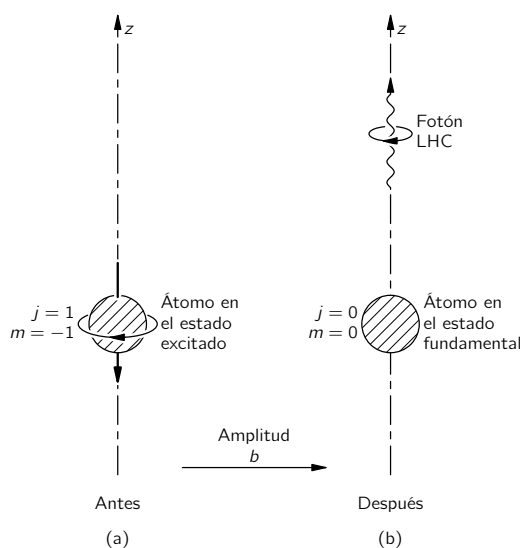


Figura 18-2. Un átomo con $m = -1$ emite un fotón LHC en dirección del eje $+z$.

- 18-1 Radiación dipolar eléctrica
- 18-2 Dispersión de la luz
- 18-3 La aniquilación del positronio
- 18-4 Matriz de rotación para un espín cualquiera
- 18-5 Midiendo un espín nuclear
- 18-6 Composición de momentum angular
- 18-7 Nota adicional 1: obtención de la matriz de rotación
- 18-8 Nota adicional 2: conservación de la paridad en la emisión de fotones

átomo queda con momentum angular cero en dirección del eje z , ya que hemos supuesto un átomo cuyo estado inferior tiene espín cero—. Denotaremos con a a la amplitud de ese proceso. De manera más precisa, a será la amplitud para emitir un fotón dentro de cierto ángulo sólido pequeño $\Delta\Omega$ centrado en el eje z , durante un tiempo dt . Observen que la amplitud para emitir un fotón LHC en la misma dirección es cero. El momentum angular resultante en dirección del eje z sería -1 para ese fotón y cero para el átomo, con un total de -1 , lo cual no conservaría el momentum angular.

Análogamente, si el espín del átomo está inicialmente “hacia abajo” (-1 en dirección del eje z), sólo puede emitir un fotón con polarización LHC en la dirección del eje $+z$, como muestra la figura 18-2. Llamaremos b a la amplitud para este proceso —entendiendo nuevamente, la amplitud para que el fotón aparezca dentro de cierto ángulo sólido $\Delta\Omega$ —. Por otra parte, si el átomo está en el estado $m = 0$, no puede emitir un fotón en la dirección $+z$, porque un fotón sólo puede tener momentum angular $+1$ o -1 en la dirección de su movimiento.

A continuación, podemos demostrar que b está relacionada con a . Supongan que hacemos una inversión de la situación en la figura 18-1, lo cual significa que debemos imaginar cuál sería el aspecto del sistema si pasáramos cada parte del mismo a un punto equivalente del lado opuesto del origen. Esto *no* significa que debemos reflejar los vectores momentum angular, porque esos vectores son artificiales. Más bien, debemos invertir el carácter real del movimiento que correspondería a ese momentum angular. En la figura 18-3(a) y (b) mostramos cuál es el aspecto de los procesos de la figura 18-1 antes y después de la inversión respecto al centro del átomo. Observen que el sentido de rotación del átomo no cambia.† En el sistema invertido de la figura 18-3(b), tenemos un átomo con $m = +1$ que emite un fotón LHC hacia abajo.

Si ahora rotamos el sistema de la figura 18-3(b) en 180° alrededor del eje x o del eje y , éste resulta idéntico al de la figura 18-2. La combinación de inversión y rotación lleva el segundo proceso al primero. Empleando la tabla 17-2, vemos que una rotación de 180° alrededor del eje y transforma un estado $m = -1$ en un estado $m = +1$, de modo que la amplitud b debe ser igual a la amplitud a , *excepto por un posible cambio de signo debido a la inversión*. El cambio de signo en la inversión dependerá de la paridad de los estados inicial y final del átomo.

En los procesos atómicos, la paridad se conserva, por lo que la paridad de todo el sistema debe ser la misma antes y después de la emisión del fotón. Lo que ocurra dependerá de si la paridad de los estados inicial y final del átomo es par o impar —la distribución angular de la radiación será diferente para casos distintos—. Tomaremos el caso común de paridad *impar* en el estado inicial y paridad *par* en el estado final; dará lo que se denomina “radiación dipolar eléctrica”. (Si los estados inicial y final tienen la misma paridad, decimos que hay “radiación dipolar magnética”, que es del tipo de la radiación proveniente de una corriente oscilante en un circuito.) Si la paridad del estado inicial es impar, su amplitud cambia de signo en la inversión, la cual lleva al sistema de (a) a (b) en la figura 18-3. El estado final del átomo tiene paridad par, por lo que su amplitud no cambia de signo. Para que la reacción conserve la paridad, la amplitud b debe ser igual a a en magnitud, pero de signo opuesto.

Concluimos que si a es la amplitud para que un estado $m = +1$ emita un fotón hacia arriba, entonces la amplitud para las paridades supuestas de los estados inicial y final para que un estado $m = -1$ emita un fotón LHC hacia arriba es $-a$.‡

Tenemos todo lo que necesitamos saber para encontrar la amplitud para que se emita un fotón en cualquier ángulo θ respecto al eje z . Supongan que

† Cuando reemplazamos x, y, z por $-x, -y, -z$, podrían pensar que todos los vectores se invierten. Esto es verdad para vectores *polares*, como los desplazamientos y las velocidades, pero *no* para vectores *axiales*, como el momentum angular —o cualquier vector que se obtenga de un producto vectorial de dos vectores polares—. Los vectores axiales tienen las mismas componentes después de una inversión.

‡ Algunos de ustedes podrían objetar los argumentos que acabamos de hacer, basándose en que los estados finales que hemos estado considerando no tienen paridad definida. Al final de este capítulo encontrarán, en la nota adicional 18-8, otra demostración que tal vez prefieran.

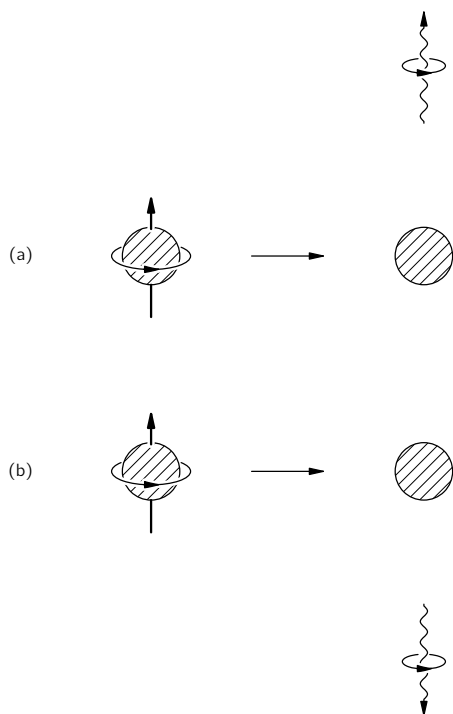


Figura 18-3. Si se transforma el proceso (a) mediante una inversión respecto al centro del átomo, aparece como en (b).

tenemos un átomo polarizado inicialmente con $m = +1$. Podemos descomponer este estado en estados $+1$, 0 y -1 respecto a un nuevo eje z' en la dirección de emisión del fotón. Las amplitudes para estos tres estados son justamente las dadas en la mitad inferior de la tabla 17-2. La amplitud para que un fotón RHC sea emitido en la dirección θ es entonces a veces la amplitud para tener $m = +1$ en esa dirección, o sea

$$a\langle + | R_y(\theta) | + \rangle = \frac{a}{2} (1 + \cos \theta). \quad (18.1)$$

La amplitud para que un fotón LHC sea emitido en la misma dirección es $-a$ veces la amplitud para tener $m = -1$ en la nueva dirección. Empleando la tabla 17-2, es

$$-a\langle - | R_y(\theta) | + \rangle = -\frac{a}{2} (1 - \cos \theta). \quad (18.2)$$

Si están interesados en otras polarizaciones, pueden encontrar las amplitudes correspondientes a partir de la superposición de estas dos amplitudes. Para obtener la intensidad de cualquier componente en función del ángulo, deben, naturalmente, tomar el cuadrado de la norma de las amplitudes.

18-2 Dispersión de la luz

Usemos estos resultados para resolver un problema algo más complicado —pero que también es más real—. Supongamos que los mismos átomos están en su estado fundamental ($j = 0$), y que *dispersan* un haz de luz incidente. Digamos que la luz va inicialmente en la dirección $+z$, de modo que tenemos fotones llegando al átomo *desde* la dirección $-z$, como muestra la figura 18-4(a). Podemos considerar a la dispersión de la luz como un proceso en dos etapas: el fotón es absorbido y luego se vuelve a emitir. Si empezamos con un fotón RHC como en la figura 18-4(a), y el momentum angular se conserva, el átomo estará en un estado $m = +1$ después de la absorción —como muestra la figura 18-4(b)—. Llamemos c a la amplitud para este proceso. Luego el átomo puede emitir un fotón RHC en la dirección θ —como en la figura 18-4(c)—. La amplitud total para que un fotón RHC sea dispersado en la dirección θ es simplemente c por (18.1). Llamemos $\langle R' | S | R \rangle$ a esta amplitud de dispersión; tenemos

$$\langle R' | S | R \rangle = \frac{ac}{2} (1 + \cos \theta). \quad (18.3)$$

También hay una amplitud para que se absorba un fotón RHC y se emita un fotón LHC. El producto de las dos amplitudes es la amplitud $\langle L' | S | R \rangle$ para que un fotón RHC sea dispersado como fotón LHC. Usando (18.2) tenemos

$$\langle L' | S | R \rangle = -\frac{ac}{2} (1 - \cos \theta). \quad (18.4)$$

Ahora veamos qué ocurre si incide un fotón LHC. Cuando se absorbe, el átomo va a un estado $m = -1$. Con el mismo razonamiento usado en la sección anterior, podemos demostrar que esta amplitud debe ser $-c$. La amplitud para que un átomo en el estado $m = -1$ emita un fotón RHC en un ángulo θ es a por la amplitud $\langle + | R_y(\theta) | - \rangle$, que es $\frac{1}{2}(1 - \cos \theta)$. Tenemos entonces

$$\langle R' | S | L \rangle = -\frac{ac}{2} (1 - \cos \theta). \quad (18.5)$$

Por último, la amplitud para que un fotón LHC sea dispersado como fotón LHC es

$$\langle L' | S | L \rangle = \frac{ac}{2} (1 + \cos \theta). \quad (18.6)$$

(Hay dos signos menos que se cancelan.)

Si hacemos una medición de la *intensidad* de dispersión para cualquier combinación de polarizaciones circulares, ésta será proporcional al cuadrado de una

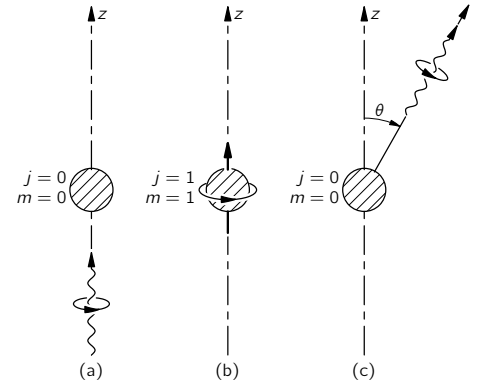


Figura 18-4. La dispersión de la luz por un átomo vista como un proceso en dos etapas.

de nuestras cuatro amplitudes. Por ejemplo, con un haz incidente de luz RHC, la intensidad de luz RHC en la radiación dispersada variará como $(1 + \cos \theta)^2$.

Todo esto está muy bien, pero supongan que comenzamos con luz *linealmente* polarizada. ¿Entonces qué pasa? Si tenemos luz polarizada en la dirección x , se la puede representar como una superposición de luz RHC y luz LHC. Escribimos (vean la sección 11-4)

$$|x\rangle = \frac{1}{\sqrt{2}}(|R\rangle + |L\rangle). \quad (18.7)$$

Y si tuviéramos luz polarizada en la dirección y , tendríamos

$$|y\rangle = -\frac{i}{\sqrt{2}}(|R\rangle - |L\rangle). \quad (18.8)$$

¿Qué quieren saber ahora? ¿Quieren la amplitud para que un fotón polarizado en la dirección x sea dispersado como un fotón RHC en un ángulo θ ? Pueden obtenerla con la regla habitual para combinar amplitudes. Primero, multiplican (18.7) por $\langle R' | S$ para obtener

$$\langle R' | S | x \rangle = \frac{1}{\sqrt{2}}(\langle R' | S | R \rangle + \langle R' | S | L \rangle), \quad (18.9)$$

y luego usan las expresiones (18.3) y (18.5) para las dos amplitudes. Obtienen

$$\langle R' | S | x \rangle = \frac{ac}{\sqrt{2}} \cos \theta. \quad (18.10)$$

Si quisieran la amplitud para que un fotón polarizado en la dirección x sea dispersado como fotón LHC, obtendrían

$$\langle L' | S | x \rangle = \frac{ac}{\sqrt{2}} \cos \theta. \quad (18.11)$$

Finalmente, supongan que quisiéramos la amplitud para que un fotón polarizado en la dirección x fuera dispersado, manteniendo su polarización en la dirección x . Lo que quieren es $\langle x' | S | x \rangle$. Esto se puede escribir en la forma

$$\langle x' | S | x \rangle = \langle x' | R' \rangle \langle R' | S | x \rangle + \langle x' | L' \rangle \langle L' | S | x \rangle. \quad (18.12)$$

Usando las relaciones

$$|R'\rangle = \frac{1}{\sqrt{2}}(|x'\rangle + i|y'\rangle), \quad (18.13)$$

$$|L'\rangle = \frac{1}{\sqrt{2}}(|x'\rangle - i|y'\rangle), \quad (18.14)$$

se obtiene que

$$\langle x' | R' \rangle = \frac{1}{\sqrt{2}}, \quad (18.15)$$

$$\langle x' | L' \rangle = \frac{1}{\sqrt{2}}. \quad (18.16)$$

Así obtienen

$$\langle x' | S | x \rangle = ac \cos \theta. \quad (18.17)$$

La respuesta es que un haz de luz polarizada en la dirección x será dispersado en la dirección θ (en el plano xz) con una *intensidad* proporcional a $\cos^2 \theta$. Si preguntan qué ocurre con la luz polarizada en la dirección y , encontrarán

$$\langle y' | S | x \rangle = 0. \quad (18.18)$$

Por lo tanto, la luz dispersada está completamente polarizada en la dirección x .

Observemos ahora algo interesante. Los resultados (18.17) y (18.18) corresponden exactamente a la teoría clásica de la dispersión de la luz que dimos en

el volumen I, sección 32-5, donde imaginamos que el electrón estaba enlazado al átomo por una fuerza restauradora lineal —de modo que actuaba como un oscilador clásico—. Quizá estén pensando: “Es mucho más fácil en la teoría clásica; si da la respuesta correcta, ¿por qué meterse en la teoría cuántica?” Por una cosa, hasta ahora sólo hemos considerado el caso especial —aunque común— de un átomo con un estado excitado $j = 1$ y un estado fundamental $j = 0$. Si el estado excitado tuviera espín dos, obtendrían un resultado diferente. Además, no hay razón para que el modelo de un electrón atado a un resorte y arrastrado por un campo eléctrico oscilante funcione para un solo fotón. Pero hemos encontrado que en realidad sí funciona, y que la polarización y las intensidades salen bien. En cierto sentido, estamos llevando gradualmente el curso entero a la certeza real. Mientras que en el volumen I hemos hecho la teoría de índices de refracción, y de la dispersión de la luz, por medio de la teoría clásica, ahora hemos demostrado que la teoría cuántica da el mismo resultado en el caso más común. En efecto, ahora hemos tratado la polarización de la luz del cielo, por ejemplo, por medio de argumentos cuánticos, que es la única manera verdaderamente adecuada.

Por supuesto, todas las teorías clásicas que funcionan deben estar apoyadas, en última instancia, en argumentos cuánticos válidos. Naturalmente, las cosas que hemos pasado tanto tiempo explicándoles, fueron escogidas precisamente de las partes de la física clásica que siguen siendo válidas en la mecánica cuántica. Notarán que no discutimos detalladamente ningún modelo del átomo con electrones dando vueltas alrededor en órbitas. Es porque ese modelo no da resultados que estén de acuerdo con la mecánica cuántica. Pero el electrón en un resorte —que, en cierto sentido, no se parece en nada al “aspecto” que tiene un átomo— sí funciona y, por lo tanto, usamos ese modelo para la teoría de índices de refracción.

18-3 La aniquilación del positronio

A continuación tomaremos un bello ejemplo. Es muy interesante y, aunque algo complicado, esperamos que no lo sea demasiado. Nuestro ejemplo es el sistema llamado *positronio*, que es un “átomo” formado por un electrón y un positrón —un estado enlazado de un e^+ y un e^- —. Es como un átomo de hidrógeno, sólo que un positrón reemplaza al protón. Este objeto tiene —como el átomo de hidrógeno— muchos estados. También como en el hidrógeno, el estado fundamental está separado en una “estructura hiperfina”, debido a la interacción de los momentos magnéticos. Los espines del electrón y del positrón son un medio, y pueden ser paralelos o antiparalelos a cualquier eje dado. (En el estado fundamental, no hay momentum angular debido al movimiento orbital.) Hay entonces cuatro estados: tres son los subestados de un sistema de espín uno, todos con la misma energía, y el otro es un estado de espín cero, con distinta energía. Sin embargo, la separación de energía es mucho mayor que los 1 420 megaciclos de hidrógeno, porque el momento magnético del positrón es mucho más intenso —1 000 veces más— que el momento del protón.

No obstante, la diferencia más importante es que el positronio no puede durar para siempre. El positrón es la antipartícula del electrón; éstos se pueden aniquilar uno a otro. Las dos partículas desaparecen completamente, convirtiendo su energía en reposo en radiación, la cual aparece como rayos γ (fotones). En la desintegración, dos partículas con masa en reposo distinta de cero se convierten en dos o más objetos que tienen masa en reposo cero.†

Comencemos analizando la desintegración del estado con espín cero del positronio. Se desintegra en dos rayos γ , con una vida media de unos 10^{-10} segundos. Inicialmente, tenemos un positrón y un electrón juntos y con espines antiparalelos, constituyendo el sistema positronio. Después de la desintegración, hay dos fotones saliendo con momenta iguales y opuestos (figura 18-5). Los momenta deben ser iguales y opuestos, porque el momentum total después de

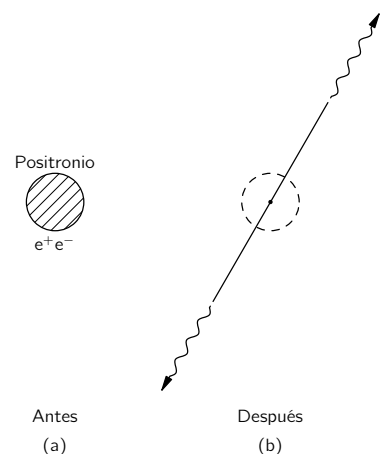


Figura 18-5. La aniquilación del positronio en dos fotones.

† En la comprensión actual más profunda del mundo, no tenemos manera fácil de distinguir si la energía de un fotón es menos “materia” que la energía de un electrón porque, como recordarán, todas las partículas se comportan de manera muy similar. La única distinción es que el fotón tiene masa en reposo cero.

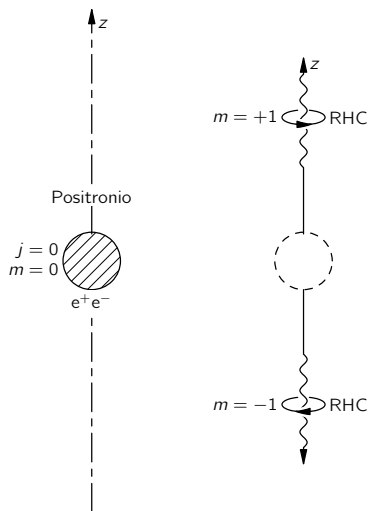


Figura 18-6. Una posibilidad de aniquilación del positronio en la dirección del eje z .

la desintegración debe ser cero, como era antes, si estamos tomando el caso de aniquilación en reposo. Si el positronio no está en reposo, podemos seguirlo, resolver el problema y luego transformar todo de vuelta al sistema de laboratorio. (Como ven, ahora podemos hacer de todo; tenemos todas las herramientas.)

En primer lugar, observamos que la distribución angular no es muy interesante. Como el estado inicial tiene espín cero, no tiene un eje especial —es simétrico frente a todas las rotaciones—. Entonces, el estado final también debe ser simétrico frente a todas las rotaciones. Esto significa que todos los ángulos de desintegración son igualmente probables —la amplitud para que el fotón vaya en cualquier dirección es la misma—. Por supuesto, una vez que encontramos a *uno* de los fotones en una dirección, el *otro* debe estar en la dirección opuesta.

La única interrogante que queda, y que queremos examinar ahora, es sobre la polarización de los fotones. Llamemos a las direcciones de movimiento de los dos fotones los ejes más y menos z . Podemos usar la representación que queremos para los estados de polarización de los fotones; para nuestra descripción, elegimos polarizaciones circulares a la derecha y a la izquierda —siempre respecto a la dirección de movimiento—.† Podemos observar inmediatamente que, si el fotón que va hacia arriba es RHC, el momentum angular se conservará si el fotón que va hacia abajo también es RHC. Cada uno llevará $+1$ unidad de momentum angular *respecto a la dirección de su momentum*, lo cual significa más y menos una unidad respecto al eje z . El total será cero, y el momentum angular después de la desintegración será el mismo que antes. Véase la figura 18-6.

El mismo argumento muestra que, si el fotón que va hacia arriba es RHC, el que va hacia abajo no puede ser LHC. Entonces, el estado final tendría dos unidades de momentum angular. Esto no está permitido si el estado inicial tiene espín cero. Observen que ese estado final tampoco es posible para el otro estado fundamental del positronio, el cual tiene espín uno, porque éste puede tener un máximo de una unidad de momentum angular en cualquier dirección.

Ahora queremos demostrar que la aniquilación en dos fotones no es posible a partir del estado de espín uno. Podrían pensar que si tomásemos el estado $j = 1$, $m = 0$ —el cual tiene momentum angular cero respecto al eje z — sería como el estado de espín cero, y se podría desintegrar en dos fotones RHC. Es claro que la desintegración mostrada en la figura 18-7(a) conserva el momentum angular

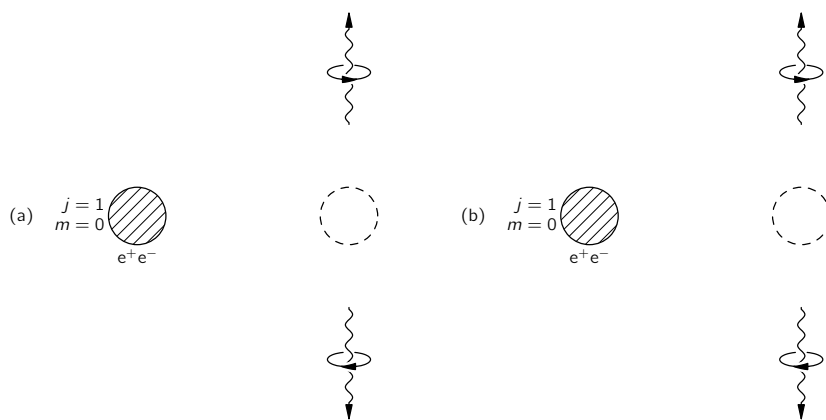


Figura 18-7. Para el estado $j = 1$ del positronio, el proceso (a) y su rotación en 180° alrededor de y (b) son exactamente iguales.

† Noten que siempre analizamos el momentum angular respecto a la dirección de movimiento de la partícula. Si nos preguntáramos acerca del momentum respecto a cualquier otro eje, tendríamos que preocuparnos de la posibilidad de momentum angular “orbital” —proveniente de un término $\mathbf{p} \times \mathbf{r}$ —. Por ejemplo, no podemos decir que los fotones salen exactamente del centro del positronio. Podrían salir disparados, como dos cosas disparadas del borde de una rueda que gira. No tenemos que preocuparnos de esas posibilidades cuando tomamos nuestro eje en la dirección de movimiento.

respecto al eje z . Pero ahora observen qué ocurre si rotamos este sistema en 180° alrededor del eje y ; obtenemos lo que muestra la figura 18-7(b). Es exactamente igual a la parte (a) de la figura. Todo lo que hemos hecho es intercambiar los dos fotones. Ahora bien, los fotones son bosones; si los intercambiamos, la amplitud tiene el mismo signo, por lo que la amplitud para la desintegración mostrada en (b) tiene que ser igual a la mostrada en (a). Pero hemos supuesto que el objeto inicial tiene un espín uno. Y cuando rotamos un objeto con espín uno en un estado $m = 0$ en 180° alrededor del eje y , sus amplitudes cambian de signo (véase la tabla 17-2 para $\theta = \pi$). Así, las amplitudes de (a) y (b) en la figura 18-7 deben tener signos opuestos; el estado de espín uno *no se puede desintegrar en dos fotones*.

Cuando se forma el positronio, sería de esperar que terminara en el estado de espín cero $1/4$ del tiempo y en el estado de espín uno (con $m = -1, 0$, o $+1$) $3/4$ del tiempo. Así, $1/4$ del tiempo tendrían aniquilaciones en dos fotones. Los otros $3/4$ del tiempo no puede haber aniquilaciones en dos fotones. Aún hay aniquilación, pero tiene que ser en *tres* fotones. Esto es más difícil para él y la vida media es 1 000 veces más larga —alrededor de 10^{-7} segundos—. Esto es lo que se observa experimentalmente. No entraremos en más detalles sobre la aniquilación del estado con espín uno.

Hasta ahora tenemos que, si nos preocupamos únicamente del momentum angular, el estado de espín cero del positronio puede convertirse en dos fotones RHC. También hay otra posibilidad: puede convertirse en dos fotones LHC, como lo muestra la figura 18-8. La pregunta que surge es, ¿cuál es la relación entre las amplitudes para estos dos modos posibles de desintegración? Podemos encontrarla a partir de la conservación de la paridad.

Sin embargo, para hacerlo, necesitamos conocer la paridad del positronio. Ahora, los físicos teóricos han demostrado, de un modo que no es fácil de explicar, que la paridad del electrón y la del positrón —su antipartícula— deben ser opuestas, por lo que el estado fundamental con espín cero del positronio debe ser impar. Supondremos simplemente que es impar y, como lo mostrarán los experimentos, podemos tomar eso como prueba suficiente.

Veamos entonces lo que ocurre si hacemos una inversión del proceso de la figura 18-6. Al hacerlo, los dos fotones invierten direcciones *y* también polarizaciones. La imagen invertida es justo como la figura 18-8. Suponiendo que la paridad del positronio es impar, las amplitudes para los dos procesos mostrados en las figuras 18-6 y 18-8 deben tener signo opuesto. Designemos con $|R_1 R_2\rangle$ al estado final de la figura 18-6, en la que ambos fotones son RHC, y denotemos con $|L_1 L_2\rangle$ al estado final de la figura 18-8, donde ambos fotones son LHC. Por lo tanto, el verdadero estado final —al que llamaremos $|F\rangle$ — debe ser

$$|F\rangle = |R_1 R_2\rangle - |L_1 L_2\rangle. \quad (18.19)$$

Entonces, una inversión intercambia las R por L y da el estado

$$P|F\rangle = |L_1 L_2\rangle - |R_1 R_2\rangle = -|F\rangle, \quad (18.20)$$

que es el negativo de (18.19). Por lo tanto, el estado final $|F\rangle$ tiene paridad negativa, la cual es igual a la del estado inicial de espín cero del positronio. Éste es el único estado final que conserva tanto el momentum angular como la paridad. Hay una amplitud para que haya desintegración en este estado, sobre lo cual, sin embargo, no tenemos por qué preocuparnos, debido a que sólo estamos interesados en cuestiones sobre la polarización.

¿Qué significa físicamente el estado final (18.19)? Uno de sus significados es el siguiente: si observamos los dos fotones en dos detectores, que se pueden hacer que cuenten separadamente fotones RHC o LHC, siempre veremos dos fotones RHC juntos o dos fotones LHC juntos. Es decir, si ustedes se paran a un lado del positronio y alguien más en el lado opuesto, ustedes pueden medir la polarización y decirle a la otra persona qué polarización obtendrá. Tienen probabilidades iguales para recibir un fotón RHC o un fotón LHC; cualquiera sea el que reciban, pueden predecir que la otra persona obtendrá lo mismo.

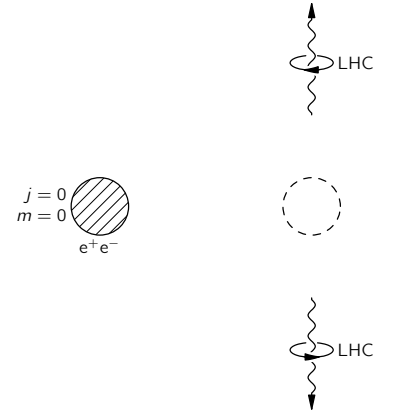


Figura 18-8. Otro proceso posible para la aniquilación del positronio.

Como hay igual probabilidad para obtener polarización RHC o LHC, parecería que hay polarización lineal. Preguntemos qué ocurre si observamos el fotón en contadores que aceptan únicamente luz polarizada linealmente. No es tan fácil medir la polarización de los rayos γ como lo es la de la luz; no hay ningún polarizador que funcione bien para longitudes de onda tan cortas. Pero imaginemos que lo hay, para facilitar la discusión. Supongan que tienen un contador que sólo acepta luz polarizada en la dirección x , y que hay un tipo del otro lado que también busca luz polarizada linealmente, digamos en la dirección y . ¿Cuál es la probabilidad para que reciban a los dos fotones provenientes de una aniquilación? Lo que necesitamos preguntar es la amplitud para que $|F\rangle$ esté en el estado $|x_1 y_2\rangle$. En otras palabras, necesitamos la amplitud

$$\langle x_1 y_2 | F \rangle,$$

que, por supuesto, es justamente

$$\langle x_1 y_2 | R_1 R_2 \rangle - \langle x_1 y_2 | L_1 L_2 \rangle. \quad (18.21)$$

Ahora bien, aunque estamos trabajando con las amplitudes de dos partículas para los dos fotones, podemos tratarlas como lo hicimos con las amplitudes para una sola partícula, ya que cada partícula actúa independientemente de la otra. Esto significa que la amplitud $\langle x_1 y_2 | R_1 R_2 \rangle$ es sencillamente el producto de las dos amplitudes independientes $\langle x_1 | R_1 \rangle$ y $\langle y_2 | R_2 \rangle$. Si usamos la tabla 17-3, entonces estas dos amplitudes son $1/\sqrt{2}$ e $i/\sqrt{2}$, de modo que

$$\langle x_1 y_2 | R_1 R_2 \rangle = +\frac{i}{2}.$$

Análogamente encontramos

$$\langle x_1 y_2 | L_1 L_2 \rangle = -\frac{i}{2}.$$

Restando estas dos amplitudes, de acuerdo con (18.21), obtenemos

$$\langle x_1 y_2 | F \rangle = +i. \quad (18.22)$$

Por lo tanto, hay probabilidad[†] *uno* de que *si* ustedes obtienen un fotón en su detector sensible a polarización en x , entonces la otra persona obtendrá un fotón en su detector sensible a polarización en y .

Supongan ahora que la otra persona ajusta su contador para polarización en la dirección x , igual que ustedes. Nunca obtendrá algún resultado cuando ustedes obtengan uno. Si lo calculan, encontrarán que

$$\langle x_1 x_2 | F \rangle = 0. \quad (18.23)$$

Por supuesto, también ocurrirá que, si ustedes ajustan su contador para polarización en la dirección y , la otra persona obtendrá resultados coincidentes, únicamente si ajusta para polarización en la dirección x .

Todo esto lleva ahora a una situación interesante. Supongan que ustedes prepararan algo como un pedazo de calcita, el cual separa los fotones en haces polarizados en las direcciones x y y , y montaran un contador en cada haz. Llamaremos contador x a uno y contador y al otro. Si quien está del otro lado hace lo mismo, siempre le pueden decir a qué haz va a ir su fotón. Siempre que ustedes y él obtienen resultados simultáneos, ustedes pueden ver cuál de los detectores recibió el fotón y decirle a la otra persona en cuál de sus detectores entró un fotón. Supongamos que, en alguna desintegración, encuentran que un fotón entró en el contador x de ustedes; pueden decirle a la otra persona que tiene que haber tenido un resultado en su contador y .

[†] No hemos normalizado nuestras amplitudes, ni las hemos multiplicado por la amplitud para la desintegración en cualquier estado final en particular, pero podemos ver que este resultado es correcto, porque obtenemos probabilidad cero cuando examinamos la otra alternativa —véase la ecuación (18.23)—.

Ahora bien, mucha gente que aprende mecánica cuántica en la forma habitual (anticuada) encuentra que esto es molesto. Les gustaría pensar que, una vez que los fotones son emitidos, todo sigue como una onda de características definidas. Pensarían que, como “cualquier fotón” tiene cierta “amplitud” para estar polarizado en la dirección x o en la dirección y , tendría que haber cierta probabilidad para recibirlo en el contador x o en el y , y que esta probabilidad no debería depender de lo que otra persona encuentre para otro fotón completamente diferente. Argumentan que “el que cualquier otra persona que haga una medición, no debería poder cambiar la probabilidad para que yo encuentre algo”. Sin embargo, nuestra mecánica cuántica dice que, al hacer una medición sobre el fotón número uno, *pueden* predecir con precisión qué polarización va a tener el fotón número dos cuando sea detectado. Este punto nunca fue aceptado por Einstein, y le preocupó mucho —se llegó a conocer como la “paradoja de Einstein-Podolsky-Rosen”—. Pero cuando se describe la situación como lo hicimos aquí, no parece que haya ninguna paradoja; se presenta en forma completamente natural que lo que se mide en un sitio está correlacionado con lo que se mide en otra parte. El argumento de que el resultado es paradójico es más o menos éste:

- (1) Si ustedes tienen un contador que les dice si su fotón es RHC o LHC, pueden predecir exactamente qué clase de fotón (RHC o LHC) encontrará la otra persona.
- (2) Por lo tanto, los fotones que la otra persona recibe deben ser sólo RHC o sólo LHC, algunos de una clase y otros de la otra.
- (3) Es seguro que no pueden alterar la naturaleza física de los fotones *de él* cambiando el tipo de observación que hacen con los fotones *de ustedes*. Cualquiera sea la medición que hagan sobre los de ustedes, los de él deben seguir siendo ya sea RHC o LHC.
- (4) Supongan ahora que la otra persona cambia su aparato para separar sus fotones en dos haces polarizados linealmente por medio de un pedazo de calcita, de modo que todos sus fotones van al haz polarizado en la dirección x o al haz polarizado en la dirección y . No hay absolutamente ninguna manera, según la mecánica cuántica, de decir a qué haz irá cualquier fotón RHC en particular. Hay 50 % de probabilidad para que vaya al haz x y 50 % de probabilidad para que vaya al haz y . Y lo mismo sucede con un fotón LHC.
- (5) Como cada fotón es RHC o LHC —de acuerdo con (2) y (3)— debe tener igual probabilidad para ir al haz x o al haz y , y no hay manera de predecir qué camino seguirá.
- (6) No obstante, la teoría predice que, si ven el fotón de *ustedes* atravesar un polarizador en dirección x , pueden predecir *con certeza* que el fotón de él irá a su haz polarizado en la dirección y . Esto está en contradicción con (5), por lo que hay una paradoja.

Sin embargo, la naturaleza aparentemente no ve la “paradoja”, porque los experimentos muestran que la predicción de (6) es, en realidad, verdadera. Ya hemos discutido la clave de esta “paradoja” en nuestra primera lección sobre el comportamiento cuántico, en el capítulo 37 del volumen I.† En la discusión anterior, los pasos (1), (2), (4) y (6) son correctos, pero (3), y su consecuencia (5), son erróneos; no son una descripción verdadera de la naturaleza. El argumento (3) dice que, por medio de *su* medición (viendo un fotón RHC o uno LHC), ustedes pueden determinar cuál de las dos alternativas se le presenta a él (ver un fotón RHC o uno LHC), *y* que aunque *no* hagan la medición, ustedes pueden decir que el proceso de éste ocurrirá de acuerdo con una de las alternativas. Pero, precisamente, en el capítulo 37 del volumen I, se insistió en señalar desde el comienzo que las cosas no son así en la naturaleza. *Su* manera requiere una

† También vean el capítulo 1 de este volumen.

descripción en términos de amplitudes que interfieren, una amplitud para cada alternativa. Una medición sobre qué alternativa se da en la realidad destruye la interferencia, pero si *no* se hace una medición, no pueden decir de todas maneras que “está ocurriendo una u otra alternativa”.

Si, para cada uno de sus fotones, pudieran determinar si éste es RHC o LHC, y *también* si está polarizado en la dirección x (todo para el mismo fotón), habría realmente una paradoja. Pero no pueden hacerlo —es un ejemplo del principio de incertidumbre—.

¿Siguen pensando que hay una “paradoja”? Asegúrense de que en efecto es una paradoja sobre el comportamiento de la naturaleza, preparando un ejemplo imaginario, en el cual la teoría de la mecánica cuántica predijera resultados incompatibles a través de dos argumentos diferentes. De otra manera, la “paradoja” es únicamente un conflicto entre la realidad y lo que ustedes sienten que la realidad “debería ser”.

¿Piensan que *no* es una “paradoja”, pero que de todas maneras es muy raro? En eso estamos de acuerdo. Es lo que hace a la física fascinante.

18-4 Matriz de rotación para un espín cualquiera

Esperamos que, a estas alturas, puedan ver lo importante que es el concepto de momentum angular en la comprensión de los procesos atómicos. Hasta ahora, sólo hemos considerado sistemas con espín —o “momentum angular total”— cero, un medio o uno. Por supuesto, hay sistemas atómicos de momentum angular mayor. Para analizar esos sistemas, necesitaríamos tablas de amplitudes para rotaciones como las de la sección 17-6. Esto es, necesitaríamos la matriz de amplitudes para espín $\frac{3}{2}$, 2, $\frac{5}{2}$, 3, etc. Aunque no elaboraremos esas tablas en detalle, queremos mostrarles cómo se hace, para que puedan hacerlo si alguna vez lo necesitan.

Como vimos antes, cualquier sistema con espín, o “momentum angular total”, j puede existir en cualquiera de los $(2j+1)$ estados, para los cuales la componente z del momentum angular puede tener alguno de los valores discretos de la sucesión $j, j-1, j-2, \dots, -(j-1), -j$ (todos en unidades de $\hbar$). Llamando $m\hbar$ a la constante z del momentum angular de cualquier estado, podemos definir un estado determinado de momentum angular dando el valor numérico de los dos “números cuánticos de momentum angular” j y m . Podemos indicar ese estado con el vector de estado $|j, m\rangle$. En el caso de una partícula con espín un medio, los dos estados son entonces $|\frac{1}{2}, \frac{1}{2}\rangle$ y $|\frac{1}{2}, -\frac{1}{2}\rangle$; y para un sistema de espín uno, en esta notación, escribiríamos $|1, +1\rangle$, $|1, 0\rangle$ y $|1, -1\rangle$. Naturalmente, una partícula con espín cero sólo tiene el estado $|0, 0\rangle$.

Ahora queremos saber qué ocurre cuando proyectamos el estado general $|j, m\rangle$ sobre una representación respecto a un sistema rotado de ejes. En primer lugar, sabemos que j es un número que caracteriza *al sistema*, por lo cual no cambia. Si rotamos los ejes, todo lo que hacemos es mezclar los diversos valores de m correspondientes al mismo j . En general, habrá cierta amplitud para que, con los ejes rotados, el sistema esté en el estado $|j, m'\rangle$, donde m' da la nueva componente z del momentum angular. Así pues, lo que queremos son todos los elementos matriciales $\langle j, m' | R | j, m \rangle$ para diversas rotaciones. Ya sabemos lo que ocurre si rotamos en un ángulo ϕ alrededor del eje z . El nuevo estado es simplemente el viejo multiplicado por $e^{im\phi}$ —aún tiene el mismo valor de m —. Podemos escribir esto así

$$R_z(\phi) |j, m\rangle = e^{im\phi} |j, m\rangle. \quad (18.24)$$

O si lo prefieren,

$$\langle j, m' | R_z(\phi) | j, m \rangle = \delta_{m, m'} e^{im\phi} \quad (18.25)$$

(donde $\delta_{m, m'}$ es 1 si $m' = m$, y cero en caso contrario).

Para una rotación alrededor de cualquier otro eje, habrá una mezcla de los diversos estados m . Naturalmente, podríamos tratar de calcular los elementos matriciales correspondientes a una rotación arbitraria descrita por los ángulos de

Euler β , α y γ . Pero es más fácil recordar que la rotación más general se puede construir con las tres rotaciones $R_z(\gamma)$, $R_y(\alpha)$ y $R_z(\beta)$; así, si conociéramos los elementos matriciales para una rotación alrededor del eje y , tendríamos todo lo que necesitamos.

¿Cómo podemos encontrar la matriz de rotación para una rotación en un ángulo θ alrededor del eje y , para una partícula con espín j ? No podemos decirles cómo hacerlo de una manera general (con lo que hemos visto). Lo hicimos para espín un medio con un complicado argumento sobre simetría. Luego, lo hicimos para espín uno, tomando el caso especial en que un sistema de espín uno estaba formado por dos partículas de espín un medio. Si continúan y aceptan el hecho de que, en el caso general, las respuestas sólo dependerán del espín j , siendo independientes de cómo se unen las entrañas del objeto con espín j , podemos extender el argumento usado para espín uno a un espín arbitrario. Por ejemplo, podemos fabricar un sistema artificial de espín $\frac{3}{2}$ con tres objetos de espín un medio. Hasta podemos evitar complicaciones imaginando que son partículas distintas —como un protón, un electrón y un muon—. Transformando cada objeto de espín un medio, podemos ver qué le ocurre al sistema total —recordando que, en el estado combinado, se multiplican las tres amplitudes—. Veamos qué ocurre en este caso.

Supongan que tomamos a los tres objetos de espín un medio con el espín “hacia arriba”; podemos indicar este estado con $|+++\rangle$. Si observamos a este sistema en un marco de referencia rotado en un ángulo ϕ alrededor del eje z , cada más sigue siendo un más, pero queda multiplicado por $e^{i\phi/2}$. Tenemos tres factores de este tipo, por lo que

$$R_z(\phi)|+++\rangle = e^{i(3\phi/2)}|+++\rangle. \quad (18.26)$$

Evidentemente, el estado $|+++\rangle$ es justamente lo que entendemos por estado $m = +\frac{3}{2}$, o estado $|\frac{3}{2}, +\frac{3}{2}\rangle$.

Si ahora rotamos este sistema alrededor del eje y , cada objeto de espín un medio tendrá cierta amplitud para ser más o para ser menos, por lo que ahora el sistema será una mezcla de las *ocho* combinaciones posibles $|+++\rangle$, $|++-\rangle$, $|+-+\rangle$, $|+--\rangle$, $|-++\rangle$, $|+-+\rangle$, $|-+-\rangle$ y $---\rangle$. Sin embargo, está claro que se las puede separar en cuatro conjuntos, que corresponden a un valor particular de m . Primero tenemos $|+++\rangle$, para el cual $m = \frac{3}{2}$. Luego están los tres estados $|++-\rangle$, $|+-+\rangle$ y $|-++\rangle$ —cada uno con dos más y un menos—. Como cada objeto con espín un medio tiene la misma probabilidad para salir con menos en la rotación, las cantidades de estas tres combinaciones deben ser iguales. Entonces tomemos la combinación

$$\frac{1}{\sqrt{3}}\{|++-\rangle + |+-+\rangle + |-++\rangle\} \quad (18.27)$$

donde se introdujo el factor $1/\sqrt{3}$ para normalizar el estado. Si rotamos este estado alrededor del eje z , obtenemos un factor $e^{i\phi/2}$ para cada más, y $e^{-i\phi/2}$ para cada menos. Cada término de (18.27) queda multiplicado por $e^{i\phi/2}$, por lo que hay un factor común $e^{i\phi/2}$. Este estado satisface nuestra idea de un estado $m = +\frac{1}{2}$; podemos concluir que

$$\frac{1}{\sqrt{3}}\{|++-\rangle + |+-+\rangle + |-++\rangle\} = |\frac{3}{2}, +\frac{1}{2}\rangle. \quad (18.28)$$

Análogamente, podemos escribir

$$\frac{1}{\sqrt{3}}\{|+- -\rangle + |-+ -\rangle + |--+\rangle\} = |\frac{3}{2}, -\frac{1}{2}\rangle, \quad (18.29)$$

que corresponde a un estado con $m = -\frac{1}{2}$. Observen que tomamos únicamente las combinaciones *simétricas* —no tomamos ninguna combinación con signos menos—. Éstas corresponderían a estados del mismo m , pero de j diferente. (Es justamente como el caso de espín uno, donde encontramos que $(1/\sqrt{2})\{|+-\rangle + |-+\rangle\}$ es

el estado $|1, 0\rangle$, pero que el estado $(1/\sqrt{2})\{|+-\rangle - |-+\rangle\}$ es el estado $|0, 0\rangle$.) Finalmente, tendríamos

$$|\frac{3}{2}, -\frac{3}{2}\rangle = |--\rangle. \quad (18.30)$$

Resumimos nuestros cuatro estados en la tabla 18-1.

Tabla 18-1

$ +++ \rangle$	$= \frac{3}{2}, +\frac{3}{2}\rangle$
$\frac{1}{\sqrt{3}}\{ ++-\rangle + +-+\rangle + -++\rangle\}$	$= \frac{3}{2}, +\frac{1}{2}\rangle$
$\frac{1}{\sqrt{3}}\{ +--\rangle + -+-\rangle + --+\rangle\}$	$= \frac{3}{2}, -\frac{1}{2}\rangle$
$--- \rangle$	$= \frac{3}{2}, -\frac{3}{2}\rangle$

Todo lo que tenemos que hacer ahora es tomar cada estado y rotarlo alrededor del eje y y ver cuánto da de los otros estados —usando nuestra matriz de rotación conocida para partículas de espín un medio—. Podemos proceder exactamente de la misma manera que en el caso de espín uno en la sección 12-6. (Sólo requiere un poco más de álgebra.) Seguiremos las ideas del capítulo 12, por lo que no repetiremos todas las explicaciones en detalle. En el sistema S , designaremos a los estados con $|\frac{3}{2}, +\frac{3}{2}, S\rangle = |+++ \rangle$, $|\frac{3}{2}, +\frac{1}{2}, S\rangle = (1/\sqrt{3})\{|++-\rangle + |+-+\rangle + |-++\rangle\}$ y así sucesivamente. El sistema T será el rotado alrededor del eje y de S en el ángulo θ . En T , los estados serán designados $|\frac{3}{2}, +\frac{3}{2}, T\rangle$, $|\frac{3}{2}, +\frac{1}{2}, T\rangle$ y así sucesivamente. Naturalmente, $|\frac{3}{2}, +\frac{3}{2}, T\rangle$ es lo mismo que $|++' +'\rangle$, donde las primas se refieren siempre al sistema T . Análogamente, $|\frac{3}{2}, +\frac{1}{2}, T\rangle$ será igual a $(1/\sqrt{3})\{|+'+'-\rangle + |+'-' +'\rangle + |-'+' +'\rangle\}$ y así sucesivamente. Cada estado $|+' \rangle$ en el sistema T proviene de los dos estados $|+\rangle$ y $|-\rangle$ en S , a través de los elementos matriciales de la tabla 12-4.

Cuando tenemos tres partículas de espín un medio, se remplaza la ecuación (12.47) por

$$\begin{aligned} |+++ \rangle &= a^3 |+'+'+' \rangle + a^2 b \{|+'+'-' \rangle + |+'-'+' \rangle + |-'+'+' \rangle\} \\ &\quad + ab^2 \{|+'-'-' \rangle + |-'+'-' \rangle + |-'-'+' \rangle\} \\ &\quad + b^3 |-'-'-' \rangle. \end{aligned} \quad (18.31)$$

Usando la transformación dada en la tabla 12-4, obtenemos en vez de (12.48) la ecuación

$$\begin{aligned} |\frac{3}{2}, +\frac{3}{2}, S\rangle &= a^3 |\frac{3}{2}, +\frac{3}{2}, T\rangle + \sqrt{3} a^2 b |\frac{3}{2}, +\frac{1}{2}, T\rangle \\ &\quad + \sqrt{3} ab^2 |\frac{3}{2}, -\frac{1}{2}, T\rangle + b^3 |\frac{3}{2}, -\frac{3}{2}, T\rangle. \end{aligned} \quad (18.32)$$

Esto ya nos da varios elementos matriciales $\langle jT | iS \rangle$. Para obtener la expresión para $|\frac{3}{2}, +\frac{1}{2}, S\rangle$, comenzamos con la transformación de un estado con dos “+” y un “-”. Por ejemplo

$$\begin{aligned} |++-\rangle &= a^2 c |+'+'+' \rangle + a^2 d |+'+'-' \rangle + abc |+'-'+' \rangle \\ &\quad + bac |-'+'+' \rangle + abd |+'-'-' \rangle + bad |-'+'-' \rangle \\ &\quad + b^2 c |-'-'+' \rangle + b^2 d |-'-'-' \rangle. \end{aligned} \quad (18.33)$$

Sumando dos expresiones análogas para $|+-+\rangle$ y $|-++\rangle$ y dividiendo entre $\sqrt{3}$, encontramos

$$\begin{aligned} |\frac{3}{2}, +\frac{1}{2}, S\rangle &= \sqrt{3} a^2 c |\frac{3}{2}, +\frac{3}{2}, T\rangle \\ &\quad + (a^2 d + 2abc) |\frac{3}{2}, +\frac{1}{2}, T\rangle \\ &\quad + (2bad + b^2 c) |\frac{3}{2}, -\frac{1}{2}, T\rangle \\ &\quad + \sqrt{3} b^2 d |\frac{3}{2}, -\frac{3}{2}, T\rangle. \end{aligned} \quad (18.34)$$

Continuando el proceso, encontramos todos los elementos $\langle jT | iS \rangle$ de la matriz de transformación que se da en la tabla 18-2. La primera columna proviene de la ecuación (18.32); la segunda de (18.34). Las dos últimas columnas se obtuvieron en la misma forma.

Tabla 18-2

Matriz de rotación para una partícula con espín $\frac{3}{2}$

(Los coeficientes a , b , c y d se dan en la tabla 12-4.)

$\langle jT iS \rangle$	$ \frac{3}{2}, +\frac{3}{2}, S\rangle$	$ \frac{3}{2}, +\frac{1}{2}, S\rangle$	$ \frac{3}{2}, -\frac{1}{2}, S\rangle$	$ \frac{3}{2}, -\frac{3}{2}, S\rangle$
$\langle \frac{3}{2}, +\frac{3}{2}, T $	a^3	$\sqrt{3}a^2c$	$\sqrt{3}ac^2$	c^3
$\langle \frac{3}{2}, +\frac{1}{2}, T $	$\sqrt{3}a^2b$	$a^2d + 2abc$	$c^2b + 2dac$	$\sqrt{3}c^2d$
$\langle \frac{3}{2}, -\frac{1}{2}, T $	$\sqrt{3}ab^2$	$2bad + b^2c$	$2cdb + d^2a$	$\sqrt{3}cd^2$
$\langle \frac{3}{2}, -\frac{3}{2}, T $	b^3	$\sqrt{3}b^2d$	$\sqrt{3}bd^2$	d^3

Ahora supongan que se rotara el sistema T respecto al S en un ángulo θ alrededor de sus ejes y . Entonces a , b , c y d tienen los valores [véase la ecuación (12.54)] $a = d = \cos \theta/2$ y $c = -b = \sin \theta/2$. Si usamos estos valores en la tabla 18-2, entonces obtenemos las expresiones que corresponden a la segunda parte de la tabla 17-2, pero ahora para un sistema de espín $\frac{3}{2}$.

Los argumentos que acabamos de dar se generalizan fácilmente a un sistema de espín j cualquiera. Además, los estados $|j, m\rangle$ se pueden armar a partir de $2j$ partículas de espín un medio. (Hay $j + m$ de ellas en el estado $|+\rangle$ y $j - m$ en el estado $|-\rangle$.) Las sumas se hacen sobre todas las maneras posibles en que se puede hacer esto, y se normaliza el estado multiplicando por una constante apropiada. Los que tengan inclinación por las matemáticas, pueden ser capaces de mostrar que se obtiene el siguiente resultado:†

$$\langle j, m' | R_y(\theta) | j, m \rangle = [(j+m)!(j-m)!(j+m')!(j-m')!]^{1/2} \times \sum_k \frac{(-1)^{k+m-m'} (\cos \theta/2)^{2j+m'-m-2k} (\sin \theta/2)^{m-m'+2k}}{(m-m'+k)!(j+m'-k)!(j-m-k)!k!}, \quad (18.35)$$

donde k toma todos los valores que dan términos ≥ 0 en todos los factoriales.

Es una fórmula bastante confusa, pero con ella pueden verificar la tabla 17-2 para $j = 1$ y preparar por su cuenta tablas para j más altos. Varios elementos matriciales especiales tienen importancia adicional y se les han dado nombres especiales. Por ejemplo, los elementos matriciales para $m = m' = 0$ y j entero se conocen como polinomios de Legendre, y se los designa $P_j(\cos \theta)$:

$$\langle j, 0 | R_y(\theta) | j, 0 \rangle = P_j(\cos \theta). \quad (18.36)$$

Los primeros polinomios de Legendre son:

$$P_0(\cos \theta) = 1, \quad (18.37)$$

$$P_1(\cos \theta) = \cos \theta, \quad (18.38)$$

$$P_2(\cos \theta) = \frac{1}{2}(3 \cos^2 \theta - 1), \quad (18.39)$$

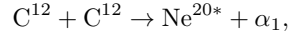
$$P_3(\cos \theta) = \frac{1}{2}(5 \cos^3 \theta - 3 \cos \theta). \quad (18.40)$$

18-5 Midiendo un espín nuclear

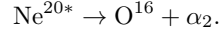
Nos gustaría mostrarles un ejemplo de la aplicación de los coeficientes que acabamos de describir. Tiene que ver con un interesante experimento reciente, el

† Si quieren detalles, pueden encontrarlos en un apéndice de este capítulo.

cual ahora serán capaces de comprender. Algunos físicos querían encontrar el espín de cierto estado excitado del núcleo de Ne^{20} . Para hacerlo, bombardearon una lámina de carbono con un haz de iones de carbono acelerados, y produjeron el estado excitado deseado del Ne^{20} —llamado Ne^{20*} — en la reacción



donde α_1 es la partícula α o He^4 . Varios estados excitados del Ne^{20} producidos de esta manera son inestables y se desintegran conforme a la reacción



Por lo que, experimentalmente, hay dos partículas α que salen de la reacción. Las llamamos α_1 y α_2 ; debido a que salen con energías diferentes, es posible distinguirlas. Además, eligiendo una energía determinada para α_1 , podemos seleccionar cualquier estado excitado en particular del Ne^{20} .

Se preparó el experimento como muestra la figura 18-9. Se hizo incidir un haz de iones de carbono de 16 MeV sobre una lámina delgada de carbono. Se contó la primera partícula α en un detector de juntura de silicio difundido, indicado con α_1 —ajustado para aceptar partículas α de la energía apropiada en dirección hacia adelante (respecto al haz incidente de C^{12})—. La segunda partícula α fue recibida en el contador α_2 en un ángulo θ respecto de α_1 . Se midió el número de señales coincidentes por unidad de tiempo provenientes de α_1 y α_2 en función del ángulo θ .

La idea del experimento es la siguiente. Primero, necesitan saber que el espín del C^{12} , del O^{16} y el de la partícula α son cero. Si llamamos dirección $+z$ a la dirección de movimiento del C^{12} inicial, entonces sabemos que el Ne^{20*} debe tener momentum angular cero respecto al eje z . Ninguna de las otras partículas tiene espín; el C^{12} llega en la dirección del eje z , y la α_1 sale en la dirección del eje z , por lo que no pueden tener momentum angular respecto a ese eje. Así, cualquiera que sea el espín j del Ne^{20*} , sabemos que está en el estado $|j, 0\rangle$. Ahora bien, ¿qué ocurrirá cuando el Ne^{20*} se desintegre en un O^{16} y la segunda partícula α ? Bueno, se recibe a la partícula α en el contador α_2 y, para que se conserve el momentum, el O^{16} debe salir en la dirección opuesta.† Respecto al nuevo eje que pasa por α_2 , no puede haber ninguna componente del momentum angular. El estado final tiene momentum angular cero respecto al nuevo eje, por lo que el Ne^{20*} se puede desintegrar en esta forma, únicamente si tiene cierta amplitud para tener m' igual a cero, donde m' es el número cuántico de la componente del momentum angular respecto al nuevo eje. De hecho, la probabilidad para observar a α_2 en el ángulo θ es el cuadrado de la amplitud (o elemento matricial)

$$\langle j, 0 | R_y(\theta) | j, 0 \rangle. \quad (18.41)$$

Para encontrar el espín del estado Ne^{20*} en cuestión, se graficó la intensidad de la segunda partícula α en función del ángulo, y se comparó con las curvas teóricas para varios valores de j . Como dijimos en la sección anterior, las amplitudes $\langle j, 0 | R_y(\theta) | j, 0 \rangle$ son simplemente las funciones $P_j(\cos \theta)$. Por lo tanto, las posibles distribuciones angulares son curvas de $[P_j(\cos \theta)]^2$. La figura 18-10 muestra los resultados experimentales para dos de los estados excitados. Pueden ver que la distribución angular para el estado de 5.80 MeV se ajusta muy bien con la curva de $[P_1(\cos \theta)]^2$, por lo que debe ser un estado de espín uno. Por el contrario, los datos para el estado de 5.63 MeV son completamente diferentes; se ajustan a la curva $[P_3(\cos \theta)]^2$. El estado tiene espín 3.

Con este experimento, hemos podido encontrar el momentum angular de dos estados excitados del Ne^{20*} . Se puede usar esta información para tratar de comprender cuál es la configuración de los protones y los neutrones dentro de este núcleo —otra pieza de información sobre las misteriosas fuerzas nucleares—.

† Podemos despreciar el retroceso impartido al Ne^{20*} en la primera colisión. O mejor aún, podemos calcularlo y hacer las correcciones correspondientes.

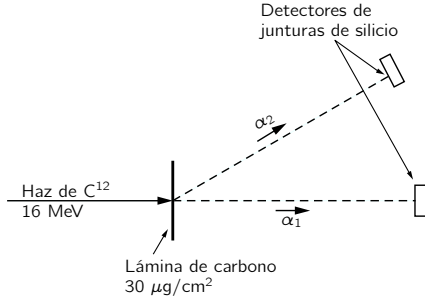


Figura 18-9. Arreglo experimental para determinar el espín de ciertos estados del Ne^{20} .

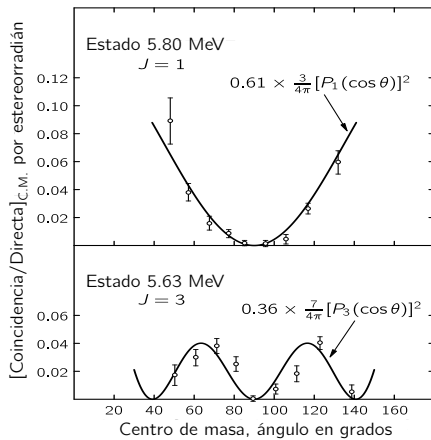


Figura 18-10. Resultados experimentales para la distribución angular de las partículas α para dos estados excitados del Ne^{20} producidos en el arreglo de la figura 18-9. [Tomada de J. A. Kuehner, *Physical Review*, vol. 125, p. 1650 (1962).]

18-6 Composición de momentum angular

Cuando estudiamos la estructura hiperfina del átomo de hidrógeno en el capítulo 12, tuvimos que calcular los estados internos de un sistema compuesto por dos partículas —el electrón y el protón—, ambas con espín un medio. Encontramos que los cuatro estados posibles de espín se podían separar en dos grupos —un grupo con una sola energía, que se veía en el mundo externo como una partícula de espín uno, y el estado restante, que se comportaba como una partícula de espín cero—. Es decir, juntando dos partículas con espín un medio, podemos formar un sistema cuyo “espín total” es uno o cero. En esta sección nos gustaría discutir, en términos más generales, los estados de espín de un *sistema* formado por dos partículas con espín arbitrario. Éste es otro problema importante sobre el momentum angular en sistemas cuánticos.

Primero, volvamos a escribir los resultados del capítulo 12 para el átomo de hidrógeno en una forma que sea más fácil de extender al caso más general. Comenzamos con dos partículas, a las que llamaremos partícula a (electrón) y partícula b (protón). La partícula a tenía espín $j_a (= \frac{1}{2})$ y la componente z de su momentum angular, m_a , podía tener diversos valores (en realidad 2, $m_a = +\frac{1}{2}$ o $m_a = -\frac{1}{2}$). Análogamente, el estado de espín para la partícula b se describe mediante su espín j_b y la componente z de su momentum angular, m_b . Pudimos formar varias combinaciones con los estados de espín de las dos partículas. Por ejemplo, podíamos tener a la partícula a con $m_a = \frac{1}{2}$ y la partícula b con $m_b = -\frac{1}{2}$, formando el estado $|a, +\frac{1}{2}; b, -\frac{1}{2}\rangle$. En general, los estados combinados formaban un sistema cuyo “espín del sistema” o “espín total” o “momentum angular total” J podía ser 1 o 0. Y el sistema podía tener una componente z del momentum angular, M , que era +1, 0 o -1 cuando $J = 1$, o 0 cuando $J = 0$. En este nuevo lenguaje, podemos reescribir las fórmulas (12.41) y (12.42) como se muestra en la tabla 18-3.

Tabla 18-3

Composición de los momentos angulares para dos partículas con espín $\frac{1}{2}$ ($j_a = \frac{1}{2}, j_b = \frac{1}{2}$).

$ J = 1, M = +1\rangle = a, +\frac{1}{2}; b, +\frac{1}{2}\rangle$
$ J = 1, M = 0\rangle = \frac{1}{\sqrt{2}} \{ a, +\frac{1}{2}; b, -\frac{1}{2}\rangle + a, -\frac{1}{2}; b, +\frac{1}{2}\rangle \}$
$ J = 1, M = -1\rangle = a, -\frac{1}{2}; b, -\frac{1}{2}\rangle$
$ J = 0, M = 0\rangle = \frac{1}{\sqrt{2}} \{ a, +\frac{1}{2}; b, -\frac{1}{2}\rangle - a, -\frac{1}{2}; b, +\frac{1}{2}\rangle \}$

La columna izquierda de la tabla describe el estado compuesto mediante su momentum angular total J y la componente z , M . La columna de la derecha muestra cómo se construyen estos estados en términos de los valores de m para las dos partículas a y b .

Ahora queremos generalizar este resultado a estados formados por dos objetos a y b con espines arbitrarios j_a y j_b . Comencemos considerando un ejemplo en el cual $j_a = \frac{1}{2}$ y $j_b = 1$, es decir, el átomo de deuterio, en el que la partícula a es un electrón (e) y la partícula b es el núcleo —un deuterón (d)—. Tenemos entonces que $j_a = j_e = \frac{1}{2}$. El deuterón está formado por un protón y un neutrón, en un estado cuyo espín total es uno, por lo que $j_b = j_d = 1$. Ahora queremos estudiar los estados hiperfinos del deuterio —como lo hicimos para el hidrógeno—. Como el deuterón tiene tres estados posibles $m_b = m_d = +1, 0, -1$, y el electrón tiene dos, $m_a = m_e = +\frac{1}{2}, -\frac{1}{2}$, están los *seis* estados posibles siguientes (usando la notación $|e, m_e; d, m_d\rangle$):

$$|e, +\frac{1}{2}; d, +1\rangle,$$

$$\begin{aligned}
&|e, +\frac{1}{2}; d, 0\rangle; |e, -\frac{1}{2}; d, +1\rangle, \\
&|e, +\frac{1}{2}; d, -1\rangle; |e, -\frac{1}{2}; d, 0\rangle, \\
&|e, -\frac{1}{2}; d, -1\rangle.
\end{aligned} \tag{18.42}$$

Notarán que hemos agrupado los estados conforme a los valores de la suma de m_e y m_d —ordenados en forma decreciente—.

Ahora nos preguntamos: ¿qué ocurre con estos estados si los proyectamos sobre un sistema de coordenadas diferente? Si simplemente se rota el nuevo sistema alrededor del eje z en un ángulo ϕ , el estado $|e, m_e; d, m_d\rangle$ queda multiplicado por

$$e^{im_e\phi} e^{im_d\phi} = e^{i(m_e+m_d)\phi}. \tag{18.43}$$

(Se puede considerar que el estado es el producto $|e, m_e\rangle |d, m_d\rangle$, y cada vector de estado contribuye independientemente con su propio factor exponencial.) El factor (18.43) es de la forma $e^{iM\phi}$, por lo que el estado $|e, m_e; d, m_d\rangle$ tiene una componente z del momentum angular igual a

$$M = m_e + m_d. \tag{18.44}$$

La componente z del momentum angular total es la suma de las componentes z del momentum angular de las partes.

Por lo tanto, en la lista dada en (18.42), el estado de la línea superior tiene $M = +\frac{3}{2}$, los dos de la segunda línea tienen $M = +\frac{1}{2}$, los dos siguientes tienen $M = -\frac{1}{2}$ y el último estado tiene $M = -\frac{3}{2}$. Vemos inmediatamente una posibilidad para el espín J de que el estado combinado (el momentum angular total) debe ser $\frac{3}{2}$, y esto requiere cuatro estados con $M = +\frac{3}{2}, +\frac{1}{2}, -\frac{1}{2}$ y $-\frac{3}{2}$.

Hay un solo candidato para $M = +\frac{3}{2}$, así que ya sabemos que

$$|J = \frac{3}{2}, M = +\frac{3}{2}\rangle = |e, +\frac{1}{2}; d, +1\rangle. \tag{18.45}$$

¿Pero cuál es el estado $|J = \frac{3}{2}, M = +\frac{1}{2}\rangle$? Tenemos dos candidatos en la segunda línea de (18.42) y, en realidad, cualquier combinación lineal de ellos también tendría $M = \frac{1}{2}$. Así, entonces, debemos esperar encontrar

$$|J = \frac{3}{2}, M = +\frac{1}{2}\rangle = \alpha |e, +\frac{1}{2}; d, 0\rangle + \beta |e, -\frac{1}{2}; d, +1\rangle, \tag{18.46}$$

donde α y β son dos números. Se llaman *coeficientes de Clebsch-Gordan*. Nuestro problema siguiente es descubrir qué son.

Podemos descubrir con facilidad sólo recordando que el deuterón está formado por un neutrón y un protón, y escribiendo los estados de deuterón más explícitamente con las reglas de la tabla 18-3. Si lo hacemos así, los estados (18.42) tienen el aspecto que muestra la tabla 18-4.

Queremos formar los cuatro estados de $J = \frac{3}{2}$ empleando los estados de la tabla. Pero ya sabemos la respuesta, porque en la tabla 18-1 tenemos estados de espín $\frac{3}{2}$ formados a partir de tres partículas con espín un medio. El primer estado de la tabla 18-1 tiene $|J = \frac{3}{2}, M = +\frac{3}{2}\rangle$, y éste es $|+++\rangle$, el cual —en la presente notación— es lo mismo que $|e, +\frac{1}{2}; n, +\frac{1}{2}; p, +\frac{1}{2}\rangle$, o sea, el primer estado de la tabla 18-4. Pero este estado también es igual al primero de la lista (18.42), lo cual confirma nuestra afirmación (18.45). La segunda línea de la tabla 18-1 dice —cambiando a nuestra notación actual— que

$$\begin{aligned}
|J = \frac{3}{2}, M = +\frac{1}{2}\rangle &= \frac{1}{\sqrt{3}} \{ |e, +\frac{1}{2}; n, +\frac{1}{2}; p, -\frac{1}{2}\rangle \\
&+ |e, +\frac{1}{2}; n, -\frac{1}{2}; p, +\frac{1}{2}\rangle + |e, -\frac{1}{2}; n, +\frac{1}{2}; p, +\frac{1}{2}\rangle \}.
\end{aligned} \tag{18.47}$$

El lado derecho se puede construir con las dos entradas de la segunda línea de la tabla 18-4, tomando $\sqrt{2/3}$ del primer término y $\sqrt{1/3}$ del segundo. En otras palabras, la ecuación (18.47) es equivalente a

$$|J = \frac{3}{2}, M = +\frac{1}{2}\rangle = \sqrt{2/3} |e, +\frac{1}{2}; d, 0\rangle + \sqrt{1/3} |e, -\frac{1}{2}; d, +1\rangle. \tag{18.48}$$

Tabla 18-4

Estados de momentum angular para un átomo de deuterio

$M = \frac{3}{2}$ $ e, +\frac{1}{2}; d, +1\rangle = e, +\frac{1}{2}; n, +\frac{1}{2}; p, +\frac{1}{2}\rangle$
$M = \frac{1}{2}$ $ e, +\frac{1}{2}; d, 0\rangle = \frac{1}{\sqrt{2}} \{ e, +\frac{1}{2}; n, +\frac{1}{2}; p, -\frac{1}{2}\rangle + e, +\frac{1}{2}; n, -\frac{1}{2}; p, +\frac{1}{2}\rangle \}$ $ e, -\frac{1}{2}; d, +1\rangle = e, -\frac{1}{2}; n, +\frac{1}{2}; p, +\frac{1}{2}\rangle$
$M = -\frac{1}{2}$ $ e, +\frac{1}{2}; d, -1\rangle = e, +\frac{1}{2}; n, -\frac{1}{2}; p, -\frac{1}{2}\rangle$ $ e, -\frac{1}{2}; d, 0\rangle = \frac{1}{\sqrt{2}} \{ e, -\frac{1}{2}; n, +\frac{1}{2}; p, -\frac{1}{2}\rangle + e, -\frac{1}{2}; n, -\frac{1}{2}; p, +\frac{1}{2}\rangle \}$
$M = -\frac{3}{2}$ $ e, -\frac{1}{2}; d, -1\rangle = e, -\frac{1}{2}; n, -\frac{1}{2}; p, -\frac{1}{2}\rangle$

Hemos encontrado nuestros dos coeficientes de Clebsch-Gordan α y β en la ecuación (18.46):

$$\alpha = \sqrt{2/3}, \quad \beta = \sqrt{1/3}. \quad (18.49)$$

Siguiendo el mismo procedimiento podemos encontrar

$$|J = \frac{3}{2}, M = -\frac{1}{2}\rangle = \sqrt{1/3} |e, +\frac{1}{2}; d, -1\rangle + \sqrt{2/3} |e, -\frac{1}{2}; d, 0\rangle. \quad (18.50)$$

Y también, por supuesto,

$$|J = \frac{3}{2}, M = -\frac{3}{2}\rangle = |e, -\frac{1}{2}; d, -1\rangle. \quad (18.51)$$

Éstas son las reglas para la composición de un espín 1 y un espín $\frac{1}{2}$ para hacer un total $J = \frac{3}{2}$. Hemos resumido (18.45), (18.48), (18.50) y (18.51) en la tabla 18-5.

Sin embargo, aquí tenemos únicamente cuatro estados, mientras que el sistema estamos considerando tiene seis estados posibles. De los dos estados de la segunda línea de (18.42) sólo hemos usado una combinación lineal para formar $|J = \frac{3}{2}, M = +\frac{1}{2}\rangle$. Hay otra combinación lineal ortogonal a la que hemos tomado, la cual también tiene $M = +\frac{1}{2}$, es decir

$$\sqrt{1/3} |e, +\frac{1}{2}; d, 0\rangle - \sqrt{2/3} |e, -\frac{1}{2}; d, +1\rangle. \quad (18.52)$$

Análogamente, se pueden combinar los dos estados que están en la tercera línea de (18.42) para dar dos estados ortogonales, cada uno con $M = -\frac{1}{2}$. El ortogonal a (18.52) es

$$\sqrt{2/3} |e, +\frac{1}{2}; d, -1\rangle - \sqrt{1/3} |e, -\frac{1}{2}; d, 0\rangle. \quad (18.53)$$

Tabla 18-5

Los estados $J = \frac{3}{2}$ del átomo de deuterio

$ J = \frac{3}{2}, M = +\frac{3}{2}\rangle = e, +\frac{1}{2}; d, +1\rangle$
$ J = \frac{3}{2}, M = +\frac{1}{2}\rangle = \sqrt{2/3} e, +\frac{1}{2}; d, 0\rangle + \sqrt{1/3} e, -\frac{1}{2}; d, +1\rangle$
$ J = \frac{3}{2}, M = -\frac{1}{2}\rangle = \sqrt{1/3} e, +\frac{1}{2}; d, -1\rangle + \sqrt{2/3} e, -\frac{1}{2}; d, 0\rangle$
$ J = \frac{3}{2}, M = -\frac{3}{2}\rangle = e, -\frac{1}{2}; d, -1\rangle$

Éstos son los dos estados que quedan. Ambos tienen $M = m_e + m_d = \pm \frac{1}{2}$, y deben ser los dos estados correspondientes a $J = \frac{1}{2}$. Tenemos entonces

$$\begin{aligned} |J = \frac{1}{2}, M = +\frac{1}{2}\rangle &= \sqrt{1/3}|e, +\frac{1}{2}; d, 0\rangle - \sqrt{2/3}|e, -\frac{1}{2}; d, +1\rangle, \\ |J = \frac{1}{2}, M = -\frac{1}{2}\rangle &= \sqrt{2/3}|e, +\frac{1}{2}; d, -1\rangle - \sqrt{1/3}|e, -\frac{1}{2}; d, 0\rangle. \end{aligned} \quad (18.54)$$

Podemos verificar que estos dos estados se comportan realmente como los estados de un objeto con espín un medio, escribiendo explícitamente las partes del deuterio en función de los estados del neutrón y del protón —empleando la tabla 18-4—. El primer estado de (18.52) es

$$\begin{aligned} \sqrt{1/6}\{&|e, +\frac{1}{2}; n, +\frac{1}{2}; p, -\frac{1}{2}\rangle + |e, +\frac{1}{2}; n, -\frac{1}{2}; p, +\frac{1}{2}\rangle\} \\ &- \sqrt{2/3}|e, -\frac{1}{2}; n, +\frac{1}{2}; p, +\frac{1}{2}\rangle, \end{aligned} \quad (18.55)$$

que también se puede escribir

$$\begin{aligned} \sqrt{1/3}\{&\sqrt{1/2}\{|e, +\frac{1}{2}; n, +\frac{1}{2}; p, -\frac{1}{2}\rangle - |e, -\frac{1}{2}; n, +\frac{1}{2}; p, +\frac{1}{2}\rangle\} \\ &+ \sqrt{1/2}\{|e, +\frac{1}{2}; n, -\frac{1}{2}; p, +\frac{1}{2}\rangle - |e, -\frac{1}{2}; n, +\frac{1}{2}; p, +\frac{1}{2}\rangle\}\}. \end{aligned} \quad (18.56)$$

Examinemos ahora los términos entre las primeras llaves y consideremos a e y p juntos. Unidos forman un estado con espín cero (véase la última línea de la tabla 18-3) y no contribuyen al momentum angular. Sólo queda el neutrón, por lo que toda la *primera* llave de (18.56) se comporta, frente a rotaciones, como un neutrón, es decir, como un estado con $J = \frac{1}{2}$, $M = +\frac{1}{2}$. Si hacemos el mismo razonamiento, entonces vemos que en la *segunda* llave de (18.56), el electrón y el neutrón se acoplan para producir momentum angular cero, y queda únicamente la contribución del protón —con $m_p = \frac{1}{2}$ —. Los términos se comportan como un objeto con $J = \frac{1}{2}$, $M = +\frac{1}{2}$. Por lo que toda la expresión (18.56) se transforma como $|J = \frac{1}{2}, M = +\frac{1}{2}\rangle$, como corresponde. El estado $M = -\frac{1}{2}$, que corresponde a (18.53), se puede construir (cambiando los $+\frac{1}{2}$ apropiados por $-\frac{1}{2}$), para obtener

$$\begin{aligned} \sqrt{1/3}\{&\sqrt{1/2}\{|e, +\frac{1}{2}; n, -\frac{1}{2}; p, -\frac{1}{2}\rangle - |e, -\frac{1}{2}; n, -\frac{1}{2}; p, +\frac{1}{2}\rangle\} \\ &+ \sqrt{1/2}\{|e, +\frac{1}{2}; n, -\frac{1}{2}; p, -\frac{1}{2}\rangle - |e, -\frac{1}{2}; n, +\frac{1}{2}; p, -\frac{1}{2}\rangle\}\}. \end{aligned} \quad (18.57)$$

Pueden verificar fácilmente que éste es igual a la segunda línea de (18.54), como debería ser si los dos términos de ese par tienen que ser los dos estados de un sistema de espín un medio. Por lo tanto, nuestros resultados están confirmado. Un deuterón y un electrón pueden existir en seis estados de espín, cuatro de los cuales se comportan como los estados de un objeto de espín $\frac{3}{2}$ (tabla 18-5) y dos de los cuales se comportan como un objeto de espín un medio (18.54).

Los resultados de la tabla 18-5 y la ecuación (18.54) fueron obtenidos usando el hecho de que el deuterón está formado por un neutrón y un protón. La validez de las ecuaciones no depende de esa condición especial. Para *cualquier* objeto de espín uno unido a cualquier objeto de espín un medio, las leyes de composición (y los coeficientes) son los mismos. El conjunto de ecuaciones de la tabla 18-5 significa que, si se rotan las coordenadas alrededor de, por ejemplo, el eje y —de modo que los estados de la partícula de espín un medio y los de la partícula de espín uno cambian de acuerdo a la tabla 17-1 y a la tabla 17-2— las combinaciones lineales del lado derecho cambiarán del modo apropiado a un objeto de espín $\frac{3}{2}$. Con la misma rotación, los estados (18.54) cambiarán como los estados de un objeto de espín un medio. Los resultados dependen únicamente de las propiedades rotacionales (es decir, de los estados de espín) de las dos partículas originales, pero no del origen de los momenta angulares. Sólo hemos utilizado este hecho para obtener las fórmulas, eligiendo un caso especial

Tabla 18-6

Composición de una partícula con espín un medio ($j_a = \frac{1}{2}$)
y una partícula con espín uno ($j_b = 1$)

$ J = \frac{3}{2}, M = +\frac{3}{2}\rangle = a, +\frac{1}{2}; b, +1\rangle$
$ J = \frac{3}{2}, M = +\frac{1}{2}\rangle = \sqrt{2/3} a, +\frac{1}{2}; b, 0\rangle + \sqrt{1/3} a, -\frac{1}{2}; b, +1\rangle$
$ J = \frac{3}{2}, M = -\frac{1}{2}\rangle = \sqrt{1/3} a, +\frac{1}{2}; b, -1\rangle + \sqrt{2/3} a, -\frac{1}{2}; b, 0\rangle$
$ J = \frac{3}{2}, M = -\frac{3}{2}\rangle = a, -\frac{1}{2}; b, -1\rangle$
$ J = \frac{1}{2}, M = +\frac{1}{2}\rangle = \sqrt{1/3} a, +\frac{1}{2}; b, 0\rangle - \sqrt{2/3} a, -\frac{1}{2}; b, +1\rangle$
$ J = \frac{1}{2}, M = -\frac{1}{2}\rangle = \sqrt{2/3} a, +\frac{1}{2}; b, -1\rangle - \sqrt{1/3} a, -\frac{1}{2}; b, 0\rangle$

en que una de las partes componentes está constituida por dos partículas de espín un medio en un estado simétrico. Reunimos todos nuestros resultados en la tabla 18-6, cambiando la notación de “e” y “d” por “a” y “b” para resaltar la generalidad de las conclusiones.

Supongan que tenemos el problema general de encontrar los estados que se pueden formar combinando dos objetos de espín arbitrario. Digamos que uno tiene j_a (por lo que su componente z , m_a , recorre los $2j_a + 1$ valores entre $-j_a$ y $+j_a$) y el otro tiene j_b (con componente z , m_b , recorriendo los valores entre $-j_b$ y $+j_b$). Los estados combinados son $|a, m_a; b, m_b\rangle$, y hay $(2j_a + 1)(2j_b + 1)$ estados diferentes. Ahora bien, ¿qué estados de espín total J se puede encontrar?

La componente z del momentum angular total, M , es igual a $m_a + m_b$, y se puede hacer una lista de estados según M [como en (18.42)]. Hay una sola manera de obtener el M más alto; éste corresponde a $m_a = j_a$ y $m_b = j_b$, y es, por lo tanto, sólo $j_a + j_b$. Esto significa que el espín total J más alto es también igual a la suma $j_a + j_b$:

$$J = (M)_{\text{máx}} = j_a + j_b.$$

Para el primer valor de M menor que $(M)_{\text{máx}}$ hay dos estados (ya sea m_a o m_b es una unidad menor que su máximo). Deben contribuir con un estado al conjunto con $J = j_a + j_b$, y el que queda debe pertenecer a un nuevo conjunto con $J = j_a + j_b - 1$. El siguiente valor de M —el tercero de la lista— se puede formar de tres maneras. (De $m_a = j_a - 2$, $m_b = j_b$; de $m_a = j_a - 1$, $m_b = j_b - 1$ y de $m_a = j_a$, $m_b = j_b - 2$.) Dos corresponden a los grupos que ya comenzamos más arriba; el tercero nos dice que también hay que incluir los estados de $J = j_a + j_b - 2$. Este procedimiento continúa hasta que, en nuestra lista, llegamos a un punto en que ya no podemos bajar un escalón más en uno de los m para formar nuevos estados.

Sea j_b el menor de j_a y j_b (si son iguales, tomen cualquiera de los dos); entonces sólo son necesarios $2j_b$ valores de J —que van de 1 en 1 desde $j_a + j_b$ hasta $j_a - j_b$ —. Esto es, cuando dos objetos de espín j_a y j_b se combinan, el sistema puede tener un momentum angular total J igual a cualquiera de los siguientes valores

$$J = \begin{cases} j_a + j_b \\ j_a + j_b - 1 \\ j_a + j_b - 2 \\ \vdots \\ |j_a - j_b|. \end{cases} \quad (18.58)$$

(Escribiendo $|j_a - j_b|$ en vez de $j_a - j_b$, podemos evitar la advertencia adicional de $j_a \geq j_b$.)

Para cada uno de estos valores de J hay $2J + 1$ estados de diferente valor de M —con M yendo de $+J$ a $-J$ —. Cada uno de éstos está formado por combinaciones lineales de los estados originales $|a, m_a; b, m_b\rangle$ con factores apropiados —los

coeficientes de Clebsch-Gordan de cada término en particular—. Podemos considerar que estos coeficientes dan la “cantidad” del estado $|j_a, m_a; j_b, m_b\rangle$ que aparece en el estado $|J, M\rangle$. Así, cada coeficiente de Clebsch-Gordan tiene, si quieren, *seis* índices que identifican su posición en las fórmulas, como las de las tablas 18-3 y 18-6. Es decir, llamando $C(J, M; j_a, m_a; j_b, m_b)$ a estos coeficientes, podríamos expresar la igualdad de la segunda línea de la tabla 18-6 escribiendo

$$C(\frac{3}{2}, +\frac{1}{2}; \frac{1}{2}, +\frac{1}{2}; 1, 0) = \sqrt{2/3},$$

$$C(\frac{3}{2}, +\frac{1}{2}; \frac{1}{2}, -\frac{1}{2}; 1, +1) = \sqrt{1/3}.$$

Aquí no calcularemos los coeficientes para ningún otro caso especial.† Sin embargo, pueden encontrar tablas en muchos libros. Ustedes mismos podrían intentar algún otro caso especial. El siguiente podría ser la composición de dos partículas de espín uno. Sólo damos el resultado final en la tabla 18-7.

Estas leyes de composición de momenta angulares son muy importantes en la física de partículas —donde tienen innumerables aplicaciones—. Por desgracia, no tenemos tiempo para examinar más ejemplos aquí.

Tabla 18-7

Composición de dos partículas con espín uno ($j_a = 1, j_b = 1$)

$ J = 2, M = +2\rangle = a, +1; b, +1\rangle$ $ J = 2, M = +1\rangle = \frac{1}{\sqrt{2}} a, +1; b, 0\rangle + \frac{1}{\sqrt{2}} a, 0; b, +1\rangle$ $ J = 2, M = 0\rangle = \frac{1}{\sqrt{6}} a, +1; b, -1\rangle + \frac{1}{\sqrt{6}} a, -1; b, +1\rangle + \frac{2}{\sqrt{6}} a, 0; b, 0\rangle$ $ J = 2, M = -1\rangle = \frac{1}{\sqrt{2}} a, 0; b, -1\rangle + \frac{1}{\sqrt{2}} a, -1; b, 0\rangle$ $ J = 2, M = -2\rangle = a, -1; b, -1\rangle$
$ J = 1, M = +1\rangle = \frac{1}{\sqrt{2}} a, +1; b, 0\rangle - \frac{1}{\sqrt{2}} a, 0; b, +1\rangle$ $ J = 1, M = 0\rangle = \frac{1}{\sqrt{2}} a, +1; b, -1\rangle - \frac{1}{\sqrt{2}} a, -1; b, +1\rangle$ $ J = 1, M = -1\rangle = \frac{1}{\sqrt{2}} a, 0; b, -1\rangle - \frac{1}{\sqrt{2}} a, -1; b, 0\rangle$
$ J = 0, M = 0\rangle = \frac{1}{\sqrt{3}} \{ a, +1; b, -1\rangle + a, -1; b, +1\rangle - a, 0; b, 0\rangle \}$

18-7 Nota adicional 1: obtención de la matriz de rotación‡

Para los que deseen ver los detalles, aquí obtendremos la matriz de rotación general para un sistema de espín (momentum angular total) j . En realidad no es muy importante calcular el caso general; una vez que tienen el concepto, pueden encontrar los resultados generales en tablas que aparecen en muchos libros. Por otra parte, después de haber llegado a estas alturas, puede que les guste ver que en verdad pueden comprender hasta las fórmulas más complicadas de la mecánica cuántica, tal como la ecuación (18.35), que aparecen en la descripción del momentum angular.

† Gran parte del trabajo está hecho ahora que tenemos la matriz de rotación general, ecuación (18.35).

‡ El material de este apéndice estaba incluido originariamente en el texto de la lección. Ahora creemos que no es necesario incluir un tratamiento tan detallado del caso general.

Extenderemos los argumentos de la sección 18-4 a un sistema de espín j , que consideraremos formado por $2j$ objetos de espín un medio. El estado con $m = j$ será $|++++\cdots+\rangle$ (con $2j$ signos más). Para $m = j - 1$, habrá $2j$ términos como $|++++\cdots+-\rangle$, $|+++ \cdots + - +\rangle$ y así sucesivamente. Ahora, consideremos el caso general, en el cual hay r mases y s menos —con $r + s = 2j$ —. En una rotación alrededor del eje z , cada uno de los r mases contribuye con $e^{+i\phi/2}$. El resultado es un cambio de fase de $(r/2 - s/2)\phi$. Pueden ver que

$$m = \frac{r - s}{2}. \quad (18.59)$$

Para $j = \frac{3}{2}$, cada estado de m definido debe ser la combinación lineal con los signos más para todos los estados de igual r y s —es decir, los estados correspondientes a cada arreglo posible con r mases y s menos—. Suponemos que ustedes pueden calcular que hay $(r + s)!/r!s!$ arreglos. Para normalizar cada estado, debemos dividir la suma entre la raíz cuadrada de este número. Podemos escribir

$$\left[\frac{(r + s)!}{r!s!} \right]^{-1/2} \{ \underbrace{++++\cdots+}_r \underbrace{---\cdots-}_s \} + (\text{todos los arreglos}) = |j, m\rangle \quad (18.60)$$

con

$$j = \frac{r + s}{2}, \quad m = \frac{r - s}{2}. \quad (18.61)$$

Nos ayudará en nuestro trabajo pasar a otra notación. Una vez que hemos definido los estados mediante la ecuación (18.60), los dos números r y s definen a un estado tan bien como j y m . Y evitaremos confundirnos si escribimos

$$|j, m\rangle = |^r_s\rangle, \quad (18.62)$$

donde, usando las igualdades (18.61),

$$r = j + m, \quad s = j - m.$$

A continuación, nos gustaría escribir la ecuación (18.60) en una nueva *notación especial* como

$$|j, m\rangle = |^r_s\rangle = \left[\frac{(r + s)!}{r!s!} \right]^{+1/2} \{ |+\rangle^r |-\rangle^s \}_{\text{perm}}. \quad (18.63)$$

Observen que hemos cambiado el exponente del factor que va al frente a *más* $\frac{1}{2}$. Lo hacemos porque hay justamente $N = (r + s)!/r!s!$ términos dentro de la llave. Comparando (18.63) con (18.60) queda claro que

$$\{ |+\rangle^r |-\rangle^s \}_{\text{perm}}$$

es simplemente un modo abreviado para escribir

$$\frac{\{ |++++\cdots- \rangle + \text{todos los arreglos} \}}{N},$$

donde N es el número de términos diferentes que hay dentro de la llave. La razón por la que esta notación es conveniente, es que cada vez que hacemos una rotación, todos los signos más contribuyen con el mismo factor, por lo que obtenemos ese factor a la potencia r -ésima. Análogamente, todos los términos s menos contribuyen con un factor a la potencia s -ésima, cualquiera sea el orden de los términos.

Ahora supongan que rotamos nuestro sistema en un ángulo θ alrededor del eje y . Lo que queremos es $R_y(\theta)|^r_s\rangle$. Cuando $R_y(\theta)$ opera sobre cada $|+\rangle$ nos da

$$R_y(\theta)|+\rangle = |+\rangle C + |-\rangle S, \quad (18.64)$$

donde $C = \cos \theta/2$ y $S = -\sin \theta/2$. Cuando $R_y(\theta)$ opera sobre cada $|-\rangle$ nos da

$$R_y(\theta) |-\rangle = |-\rangle C - |+\rangle S.$$

Entonces, lo que queremos es

$$\begin{aligned} R_y(\theta) \begin{bmatrix} r \\ s \end{bmatrix} &= \left[\frac{(r+s)!}{r!s!} \right]^{1/2} R_y(\theta) \{ |+\rangle^r |-\rangle^s \}_{\text{perm}} \\ &= \left[\frac{(r+s)!}{r!s!} \right]^{1/2} \{ (R_y(\theta) |+\rangle)^r (R_y(\theta) |-\rangle)^s \}_{\text{perm}} \\ &= \left[\frac{(r+s)!}{r!s!} \right]^{1/2} \{ (|+\rangle C + |-\rangle S)^r (|-\rangle C - |+\rangle S)^s \}_{\text{perm}}. \end{aligned} \quad (18.65)$$

Ahora bien, hay que desarrollar cada binomio a la potencia adecuada y multiplicar ambas expresiones. Habrá términos con $|+\rangle$ a todas las potencias, desde cero hasta $(r+s)$. Examinemos todos los términos que tienen $|+\rangle$ a la potencia r' . Siempre aparecerán multiplicados por $|-\rangle$ a la potencia s' , donde tenemos que $s' = 2j - r'$. Supongan que juntamos todos esos términos. Para cada permutación, tendrán cierto coeficiente numérico en el que intervienen los factores del binomio y los factores C y S . Supongan que llamamos $A_{r'}$ a ese factor. Entonces la ecuación (18.65) tendrá la siguiente forma

$$R_y(\theta) \begin{bmatrix} r \\ s \end{bmatrix} = \sum_{r'=0}^{r+s} \{ A_{r'} |+\rangle^{r'} |-\rangle^{s'} \}_{\text{perm}}. \quad (18.66)$$

Ahora dividamos $A_{r'}$ entre el factor $[(r'+s')!/r'!s']^{1/2}$ y llamemos $B_{r'}$ al cociente. La ecuación (18.66) es entonces equivalente a

$$R_y(\theta) \begin{bmatrix} r \\ s \end{bmatrix} = \sum_{r'=0}^{r+s} B_{r'} \left[\frac{(r'+s')!}{r'!s'!} \right]^{1/2} \{ |+\rangle^{r'} |-\rangle^{s'} \}_{\text{perm}}. \quad (18.67)$$

[Podríamos decir simplemente que esta ecuación define a $B_{r'}$ por la condición de que (18.67) dé la misma expresión que aparece en (18.65).]

Con esta definición para $B_{r'}$, los factores que quedan en el lado derecho de la ecuación (18.67) son justamente los estados $\begin{bmatrix} r' \\ s' \end{bmatrix}$. Tenemos así

$$R_y(\theta) \begin{bmatrix} r \\ s \end{bmatrix} = \sum_{r'=0}^{r+s} B_{r'} \begin{bmatrix} r' \\ s' \end{bmatrix}, \quad (18.68)$$

con s' siempre igual a $r+s-r'$. Esto significa, naturalmente, que los coeficientes $B_{r'}$ son justamente los elementos matriciales que queremos, o sea

$$\langle \begin{bmatrix} r' \\ s' \end{bmatrix} | R_y(\theta) | \begin{bmatrix} r \\ s \end{bmatrix} \rangle = B_{r'}. \quad (18.69)$$

Ahora sólo nos queda pasar por el trabajo algebraico de encontrar los diversos $B_{r'}$. Comparando (18.65) con (18.67) —y recordando que $r'+s' = r+s$ — vemos que $B_{r'}$ es simplemente el coeficiente de $a^{r'}b^{s'}$ en la siguiente expresión:

$$\left(\frac{r'!s'!}{r!s!} \right)^{1/2} (aC + bS)^r (bC - aS)^s. \quad (18.70)$$

Ahora queda únicamente el trabajo desagradable de hacer los desarrollos mediante el teorema del binomio, y juntar todos los términos con las potencias adecuadas de a y b . Si se toman el trabajo, encontrarán que el coeficiente de $a^{r'}b^{s'}$ en (18.70) es

$$\left[\frac{r'!s'!}{r!s!} \right]^{1/2} \sum_k (-1)^k S^{r-r'+2k} C^{s+r'-2k} \cdot \frac{r!}{(r-r'+k)!(r'-k)!} \cdot \frac{s!}{(s-k)!k!}. \quad (18.71)$$

La suma se debe hacer sobre todos los enteros k que dan números mayores o iguales que cero en los factoriales. Esta expresión es entonces el elemento matricial que queríamos.

Finalmente, podemos regresar a nuestra notación original en términos de j , m y m' usando

$$r = j + m, \quad r' = j + m', \quad s = j - m, \quad s' = j - m'.$$

Haciendo estas sustituciones, obtenemos la ecuación (18.35) de la sección 18-4.

18-8 Nota adicional 2: conservación de la paridad en la emisión de fotones

En la sección 18-1 de este capítulo, consideramos la emisión de luz por un átomo que pasa de un estado excitado con espín 1 a un estado fundamental con espín 0. Si el estado excitado tiene el espín hacia arriba ($m = +1$), puede emitir un fotón RHC en el eje $+z$ o un fotón LHC en el eje $-z$. Llamemos a estos dos estados del fotón $|R_{ar}\rangle$ y $|L_{ab}\rangle$. Ninguno de estos estados tiene paridad definida. Sea $\hat{P}$ el operador de paridad, $\hat{P}|R_{ar}\rangle = |L_{ab}\rangle$ y $\hat{P}|L_{ab}\rangle = |R_{ar}\rangle$.

¿Y entonces qué pasa con nuestra demostración anterior de que un átomo en un estado de energía definida debe tener paridad definida, y con nuestra afirmación de que la paridad se conserva en los procesos atómicos? El estado final en este problema (el estado después de la emisión de un fotón), ¿no debería tener paridad definida? *Sí* la tiene si consideramos el estado final *completo*, el cual contiene amplitudes para la emisión de fotones en todos los ángulos. En la sección 18-1 escogimos considerar sólo una parte del estado final completo.

Si lo deseamos, podemos considerar sólo los estados finales que sí tienen paridad definida. Por ejemplo, consideren un estado final $|\psi_F\rangle$, el cual tiene cierta amplitud α para ser un fotón RHC que va en la dirección $+z$, y cierta amplitud β de ser un fotón LHC que va en la dirección $-z$. Podemos escribir

$$|\psi_F\rangle = \alpha |R_{ar}\rangle + \beta |L_{ab}\rangle. \quad (18.72)$$

La operación de paridad sobre este estado da

$$\hat{P}|\psi_F\rangle = \alpha |L_{ab}\rangle + \beta |R_{ar}\rangle. \quad (18.73)$$

Este estado será $\pm|\psi_F\rangle$ si $\beta = \alpha$ o si $\beta = -\alpha$. Por lo tanto, un estado final de paridad par es

$$|\psi_F^+\rangle = \alpha\{|R_{ar}\rangle + |L_{ab}\rangle\}, \quad (18.74)$$

y un estado de paridad impar es

$$|\psi_F^-\rangle = \alpha\{|R_{ar}\rangle - |L_{ab}\rangle\}. \quad (18.75)$$

A continuación, queremos considerar la desintegración de un estado excitado de paridad impar hacia un estado fundamental de paridad par. Si se tiene que conservar la paridad, el estado final del fotón debe tener paridad impar. Debe ser el estado en (18.75). Si la amplitud para encontrar $|R_{ar}\rangle$ es α , la amplitud para encontrar $|L_{ab}\rangle$ es $-\alpha$.

Observen ahora lo que ocurre cuando hacemos una rotación de 180° alrededor del eje y . El estado excitado inicial del átomo se convierte en un estado $m = -1$ (sin cambio de signo, de acuerdo con la tabla 17-2). Y la rotación del estado final da

$$R_y(180^\circ)|\psi_F^-\rangle = \alpha\{|R_{ab}\rangle - |L_{ar}\rangle\}. \quad (18.76)$$

Comparando esta ecuación con (18.75) ven que, para la paridad supuesta del estado final, la amplitud para obtener un fotón LHC en la dirección $+z$ a partir del estado inicial $m = -1$ es el negativo de la amplitud para obtener un fotón RHC a partir del estado inicial $m = +1$. Esto concuerda con los resultados encontrados en la sección 18-1.

El átomo de hidrógeno y la tabla periódica

19-1 La ecuación de Schrödinger para el átomo de hidrógeno

El éxito más significativo en la historia de la mecánica cuántica fue la comprensión de los detalles de los espectros de algunos átomos simples y la comprensión de las periodicidades que se encuentran en la tabla de los elementos químicos. En este capítulo llevaremos, por fin, nuestra mecánica cuántica al tema de este hecho importante, específicamente a la comprensión del espectro del átomo de hidrógeno. Al mismo tiempo llegaremos a una explicación cualitativa de las propiedades misteriosas de los elementos químicos. Conseguiremos esto al estudiar en detalle el comportamiento del electrón en el átomo de hidrógeno —y haremos por primera vez un cálculo detallado de una distribución espacial de acuerdo a los conceptos que desarrollamos en el capítulo 16—.

Para una descripción completa del átomo de hidrógeno deberíamos describir el movimiento tanto del protón como del electrón. Es posible hacer esto en la mecánica cuántica de un modo análogo a como se hace en la mecánica clásica: se describe el movimiento de cada partícula relativo al centro de masa, pero no lo haremos así. Solamente discutiremos una aproximación en la que consideraremos que el protón es muy pesado y, por lo tanto, podemos considerarlo como fijo en el centro del átomo.

Haremos otra aproximación olvidándonos de que el electrón tiene espín y debería ser descrito mediante leyes relativistas de la mecánica. Serán necesarias algunas pequeñas correcciones más, puesto que usaremos la ecuación de Schrödinger no relativista y despreciaremos los efectos magnéticos. Hay efectos magnéticos pequeños ya que desde el punto de vista del electrón, el protón es una carga en movimiento que produce un campo magnético. En este campo el electrón tendrá una energía diferente con un espín hacia arriba que con él hacia abajo. La energía del átomo estará desviada un poquito respecto a lo que nosotros calcularemos. Ignoraremos este pequeño corrimiento de energía. Además imaginaremos que el electrón es exactamente como un giroscopio que se mueve en el espacio conservando siempre la misma dirección de espín. Como estaremos considerando un átomo libre en el espacio, el momentum angular total se conservará. Supondremos en nuestra aproximación que el momentum angular del espín del electrón permanece constante, por lo que el resto del momentum angular del átomo —que se llama corrientemente momentum angular “orbital”— también se conservará. En una excelente aproximación, el electrón se mueve en el átomo de hidrógeno como una partícula sin espín —el momentum angular del movimiento es una constante—.

Con estas aproximaciones, la amplitud de encontrar el electrón en lugares diferentes del espacio se puede representar por una función de la posición en el espacio y en el tiempo. Sea $\psi(x, y, z, t)$ la amplitud de encontrar el electrón en alguna parte en el tiempo t . Según la mecánica cuántica, la derivada de esta amplitud respecto al tiempo está dada por el operador hamiltoniano aplicado a la misma función. Según el capítulo 16,

$$i\hbar \frac{\partial \psi}{\partial t} = \hat{\mathcal{H}}\psi, \quad (19.1)$$

con

$$\hat{\mathcal{H}} = -\frac{\hbar^2}{2m} \nabla^2 + V(\mathbf{r}). \quad (19.2)$$

19-1 La ecuación de Schrödinger para el átomo de hidrógeno

19-2 Soluciones con simetría esférica

19-3 Estados con dependencia angular

19-4 Solución general para el hidrógeno

19-5 Funciones de onda del hidrógeno

19-6 La tabla periódica

Aquí, m es la masa del electrón y $V(\mathbf{r})$ es la energía potencial del electrón en el campo electrostático del protón. Tomando $V = 0$ a grandes distancias del protón, podemos escribir†

$$V = -\frac{e^2}{r}.$$

La función de onda ψ debe, pues, satisfacer la ecuación

$$i\hbar \frac{\partial \psi}{\partial t} = -\frac{\hbar^2}{2m} \nabla^2 \psi - \frac{e^2}{r} \psi. \quad (19.3)$$

Queremos buscar estados de energía definida, por lo cual buscamos soluciones que tengan la forma

$$\psi(\mathbf{r}, t) = e^{-(i/\hbar)Et} \psi(\mathbf{r}). \quad (19.4)$$

La función $\psi(\mathbf{r})$ debe entonces ser una solución de

$$-\frac{\hbar^2}{2m} \nabla^2 \psi = \left(E + \frac{e^2}{r}\right) \psi, \quad (19.5)$$

donde E es alguna constante —la energía del átomo—.

Como la energía potencial depende solamente del radio, resulta que es mucho más conveniente resolver esta ecuación en coordenadas polares que en coordenadas cartesianas. El laplaciano se define en coordenadas cartesianas mediante:

$$\nabla^2 = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}.$$

Queremos usar en su lugar las coordenadas r, θ, ϕ que muestra la figura 19-1. Estas coordenadas se relacionan con x, y, z en la forma:

$$x = r \sin \theta \cos \phi; \quad y = r \sin \theta \sin \phi; \quad z = r \cos \theta.$$

Es una cosa tediosa calcularlo, pero ustedes pueden eventualmente demostrar que para cualquier función $f(\mathbf{r}) = f(r, \theta, \phi)$,

$$\nabla^2 f(r, \theta, \phi) = \frac{1}{r} \frac{\partial^2}{\partial r^2} (rf) + \frac{1}{r^2} \left\{ \frac{1}{\sin \theta} \frac{\partial}{\partial \theta} \left(\sin \theta \frac{\partial f}{\partial \theta} \right) + \frac{1}{\sin^2 \theta} \frac{\partial^2 f}{\partial \phi^2} \right\}. \quad (19.6)$$

Así, en coordenadas polares la ecuación que $\psi(r, \theta, \phi)$ tiene que satisfacer:

$$\frac{1}{r} \frac{\partial^2}{\partial r^2} (r\psi) + \frac{1}{r^2} \left\{ \frac{1}{\sin \theta} \frac{\partial}{\partial \theta} \left(\sin \theta \frac{\partial \psi}{\partial \theta} \right) + \frac{1}{\sin^2 \theta} \frac{\partial^2 \psi}{\partial \phi^2} \right\} = -\frac{2m}{\hbar^2} \left(E + \frac{e^2}{r} \right) \psi. \quad (19.7)$$

19-2 Soluciones con simetría esférica

Trataremos en primer lugar de encontrar cierta función muy simple que satisfaga la horrible ecuación (19.7). Aunque la función de onda ψ dependerá, en general, de los ángulos θ y ϕ así como el radio r , podemos ver si podría haber una situación especial en que ψ *no* dependa de los ángulos. Para una función de onda que no depende de los ángulos, ninguna de las amplitudes cambiará si se rota el sistema de coordenadas. Esto significa que todas las componentes del momentum angular son cero. Tal ψ debe corresponder a un estado cuyo momentum angular total es cero (en nuestro caso, sólo el momentum angular orbital es el que es cero, pues ignoramos el espín del electrón, aunque lo tenemos). Un estado de momentum angular orbital cero se denomina con un nombre especial. Se llama “estado s ” —pueden recordarlo: “ s significa con simetría esférica”—.‡

† Como de costumbre, $e^2 = q_e^2/4\pi\epsilon_0$.

‡ Como estos nombres especiales son parte del vocabulario corriente en la física atómica, se los tendrán que aprender. Les ayudaremos colocándolos juntos en un “diccionario” corto más adelante en este capítulo.

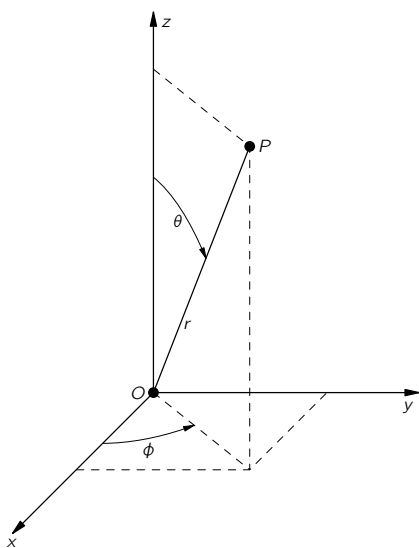


Figura 19-1. Las coordenadas polares esféricas r, θ, ϕ del punto P .

Ahora bien, si ψ no va depender de θ ni de ϕ el laplaciano completo contiene solamente el primer término y la ecuación (19.7) se torna mucho más simple:

$$\frac{1}{r} \frac{d^2}{dr^2} (r\psi) = -\frac{2m}{\hbar^2} \left(E + \frac{e^2}{r} \right) \psi. \quad (19.8)$$

Antes de que comiencen a resolver una ecuación como ésta, es bueno desembarazarse de todas las constantes adicionales tales como e^2 , m y $\hbar$ efectuando algunos cambios de escala. De este modo el álgebra será más fácil. Si hacemos las siguientes sustituciones:

$$r = \frac{\hbar^2}{me^2} \rho, \quad (19.9)$$

y

$$E = \frac{me^4}{2\hbar^2} \epsilon, \quad (19.10)$$

la ecuación (19.8) se convierte (después de multiplicar todo por ρ) en

$$\frac{d^2(\rho\psi)}{d\rho^2} = -\left(\epsilon + \frac{2}{\rho} \right) \rho\psi. \quad (19.11)$$

Los cambios de escala significan que estamos midiendo la distancia r y la energía E como múltiplos de unidades atómicas “naturales”. Esto es, $\rho = r/r_B$, donde $r_B = \hbar^2/me^2$ es el llamado “radio de Bohr” y vale alrededor de 0.528 ángstroms. Análogamente, $\epsilon = E/E_R$, con $E_R = me^4/2\hbar^2$. Ésta se llama “energía de Rydberg” y vale alrededor de 13.6 electronvoltios.

Como el producto $\rho\psi$ aparece en ambos miembros, es conveniente trabajar con él en lugar de con ψ . Poniendo

$$\rho\psi = f, \quad (19.12)$$

tenemos la ecuación de apariencia más simple

$$\frac{d^2 f}{d\rho^2} = -\left(\epsilon + \frac{2}{\rho} \right) f. \quad (19.13)$$

Ahora tenemos que hallar alguna función f , que la satisfaga la ecuación (19.13)—en otras palabras, tenemos que resolver justamente una ecuación diferencial—. Desgraciadamente no hay algún método general útil para resolver cualquier ecuación diferencial dada. No tienen más que tantear. Nuestra ecuación no es fácil, pero los físicos han encontrado que se puede resolver por el siguiente procedimiento. En primer lugar, replacen f , que es cierta función de ρ , por un producto de dos funciones

$$f(\rho) = e^{-\alpha\rho} g(\rho). \quad (19.14)$$

Esto significa simplemente que están sacando el factor común $e^{-\alpha\rho}$ de $f(\rho)$. Pueden ciertamente hacer esto para cualquier $f(\rho)$. Esto solamente desvía nuestro problema a encontrar la función $g(\rho)$ correcta.

Introduciendo (19.14) en (19.13), obtenemos la siguiente ecuación en g :

$$\frac{d^2 g}{d\rho^2} - 2\alpha \frac{dg}{d\rho} + \left(\frac{2}{\rho} + \epsilon + \alpha^2 \right) g = 0. \quad (19.15)$$

Como tenemos la libertad de escoger α , hagamos:

$$\alpha^2 = -\epsilon, \quad (19.16)$$

y obtenemos

$$\frac{d^2 g}{d\rho^2} - 2\alpha \frac{dg}{d\rho} + \frac{2}{\rho} g = 0. \quad (19.17)$$

Pueden pensar que no estamos en mejor situación que con la ecuación (19.13), pero lo interesante de la nueva ecuación es que se puede resolver fácilmente en

términos de una serie de potencias de ρ . (Es posible, en principio, resolver (19.13) también de ese modo, pero es más difícil.) Estamos diciendo que (19.17) puede ser satisfecha por cierta $g(\rho)$ que se puede escribir como una serie,

$$g(\rho) = \sum_{k=1}^{\infty} a_k \rho^k, \quad (19.18)$$

donde los a_k son coeficientes constantes. Ahora todo lo que tenemos que hacer es encontrar un apropiado conjunto infinito de coeficientes! Verifiquemos si tal solución sirve. La primera derivada de esta $g(\rho)$ es

$$\frac{dg}{d\rho} = \sum_{k=1}^{\infty} a_k k \rho^{k-1},$$

y la segunda derivada es

$$\frac{d^2g}{d\rho^2} = \sum_{k=1}^{\infty} a_k k(k-1) \rho^{k-2}.$$

Introduciendo estas expresiones en (19.17) tenemos

$$\sum_{k=1}^{\infty} k(k-1)a_k \rho^{k-2} - \sum_{k=1}^{\infty} 2\alpha k a_k \rho^{k-1} + \sum_{k=1}^{\infty} 2a_k \rho^{k-1} = 0. \quad (19.19)$$

No es evidente que hayamos tenido éxito; pero seguimos adelante. Todo parecerá mejor si reemplazamos la primera suma por un equivalente. Como el primer término de la suma es cero, podemos reemplazar cada k por $k+1$ sin cambiar nada en la serie infinita; con este cambio la primera suma se puede igualmente escribir en la forma

$$\sum_{k=1}^{\infty} (k+1)k a_{k+1} \rho^{k-1}.$$

Ahora podemos colocar todas las sumas juntas para obtener

$$\sum_{k=1}^{\infty} [(k+1)k a_{k+1} - 2\alpha k a_k + 2a_k] \rho^{k-1} = 0. \quad (19.20)$$

Esta serie de potencias se debe anular para todos los valores posibles de ρ . Esto se puede hacer solamente si el coeficiente de cada potencia de ρ es cero por separado. Tendremos una solución para el átomo de hidrógeno si podemos encontrar un conjunto de a_k para el cual

$$(k+1)k a_{k+1} - 2(\alpha k - 1)a_k = 0 \quad (19.21)$$

para toda $k \geq 1$. Esto es ciertamente fácil de ordenar. Tomen cualquier a_1 que deseen. Luego calculen todos los demás coeficientes a partir de:

$$a_{k+1} = \frac{2(\alpha k - 1)}{k(k+1)} a_k. \quad (19.22)$$

Con esto obtendrán a_2, a_3, a_4 , etc., y cada par ciertamente satisfará (19.21). Obtenemos una serie para $g(\rho)$ que satisface (19.17). Con ella podemos formar una ψ que satisfaga la ecuación de Schrödinger. Observen que las soluciones dependen de la energía supuesta (a través de α), pero para cada valor de ϵ , hay una serie correspondiente.

Tenemos una solución, pero ¿qué representa físicamente? Podemos darnos una idea viendo lo que sucede lejos del protón —para grandes valores de ρ —. Allá, los términos de la serie de mayor orden son los más importantes, por lo que

debemos considerar lo que sucede para k grandes. Cuando $k \gg 1$, la ecuación (19.22) es aproximadamente igual a

$$a_{k+1} = \frac{2\alpha}{k} a_k,$$

lo cual significa que

$$a_{k+1} \approx \frac{(2\alpha)^k}{k!}. \quad (19.23)$$

Pero éstos son justamente los coeficientes de la serie para $e^{+2\alpha\rho}$. La función de g es una exponencial que aumenta rápidamente. Aun acoplada con $e^{-\alpha\rho}$ para producir $f(\rho)$ —véase la ecuación (10.14)— da por igual una solución para $f(\rho)$ que varía como $e^{\alpha\rho}$ para ρ grandes. Encontramos una solución matemática, pero no una física. ¡Representa una situación en la cual el electrón tiene *la menor* probabilidad de estar cerca del protón! Es siempre más probable que se le encuentre para radios ρ grandes. Una función de onda para un electrón *ligado* debe tender a cero para ρ grandes.

Tenemos que pensar si hay algún modo de ganar el juego, y lo hay. ¡Observen! Si tuviésemos la suerte de que α fuese igual a $1/n$ positivo, donde n es cualquier entero, la ecuación (19.22) reportaría $a_{n+1} = 0$. Todos los términos mayores también serían cero. No tendríamos una serie infinita, sino un polinomio finito. Cualquier polinomio aumenta más lentamente que $e^{\alpha\rho}$, por lo cual eventualmente el término $e^{-\alpha\rho}$ lo vencerá y la función f tenderá a cero para ρ grandes. Las únicas soluciones para *estados ligados* son aquellas para las que $\alpha = 1/n$, con $n = 1, 2, 3, 4$, etcétera.

Volviendo a la ecuación (19.16), vemos que las soluciones para estados ligados de la ecuación de onda con simetría esférica pueden existir solamente cuando:

$$-\epsilon = 1, \frac{1}{4}, \frac{1}{9}, \frac{1}{16}, \dots, \frac{1}{n^2}, \dots$$

Las energías permitidas son solamente esas fracciones por el Rydberg, $E_R = me^4/2\hbar^2$; por lo tanto, la energía del n -ésimo nivel de energía es

$$E_n = -E_R \frac{1}{n^2}. \quad (19.24)$$

Incidentalmente, no hay nada misterioso en los números negativos para la energía. Las energías son negativas porque cuando escogimos escribir $V = -e^2/r$, tomamos como cero de energía la de un electrón colocado lejos del protón. Cuando está cerca del protón su energía es menor, es decir, algo menor que cero. El valor más bajo (más negativo) es para $n = 1$ y aumenta hacia cero con el aumento de n .

Antes del descubrimiento de la mecánica cuántica, se sabía por estudios experimentales del espectro de hidrógeno que los niveles de energía se podían describir mediante la ecuación (19.24), donde para E_R se encontró un valor de alrededor de 13.6 electronvoltios. Entonces Bohr ideó un modelo que dio la misma ecuación y predijo que E_R debería ser $me^4/2\hbar^2$. Éste fue el primer gran éxito de la teoría de Schrödinger: poder reproducir este resultado a partir de una ecuación básica de movimiento para el electrón.

Ahora que hemos resuelto nuestro primer átomo, consideremos la naturaleza de la solución que obtuvimos. Reuniendo todas sus partes, cada solución se presenta como a continuación:

$$\psi_n = \frac{f_n(\rho)}{\rho} = \frac{e^{-\rho/n}}{\rho} g_n(\rho), \quad (19.25)$$

donde

$$g_n(\rho) = \sum_{k=1}^n a_k \rho^k \quad (19.26)$$

y

$$a_{k+1} = \frac{2(k/n - 1)}{k(k+1)} a_k. \quad (19.27)$$

Ya que estamos principalmente interesados en las probabilidades relativas de encontrar el electrón en diversos lugares, podemos escoger el número que deseemos para a_1 . Podemos, por consiguiente, poner $a_1 = 1$. (Los físicos con frecuencia escogen a_1 de manera que la función de onda esté “normalizada”; es decir, de manera que la probabilidad integrada de encontrar el electrón en cualquier sitio del átomo sea igual a 1. Ahora no tenemos necesidad de hacer esto.)

Para el estado más bajo de energía, $n = 1$, y

$$\psi_1(\rho) = e^{-\rho}. \quad (19.28)$$

Para un átomo de hidrógeno en su estado fundamental (energía más baja), la amplitud de encontrar el electrón en cualquier punto decrece exponencialmente con la distancia al protón. Es más probable que se le encuentre en el protón mismo y la distancia característica hasta donde está diseminado es aproximadamente una unidad ρ , o sea, alrededor de un radio de Bohr, r_B .

Si hacemos $n = 2$ tenemos el nivel próximo más alto. La función de onda para este estado tiene dos términos. Es

$$\psi_2(\rho) = \left(1 - \frac{\rho}{2}\right) e^{-\rho/2}. \quad (19.29)$$

La función de onda para el nivel siguiente es:

$$\psi_3(\rho) = \left(1 - \frac{2\rho}{3} + \frac{2}{27}\rho^2\right) e^{-\rho/3}. \quad (19.30)$$

Las funciones de onda de estos tres primeros niveles están representadas en la figura 19-2. Pueden ver la marcha general. Todas las funciones de onda se acercan a cero rápidamente para ρ grandes después de oscilar algunas veces. De hecho, el número de “gibas” es exactamente igual a n —o, si lo prefieren, el número de veces que ψ_n pasa por cero es $n - 1$ —.

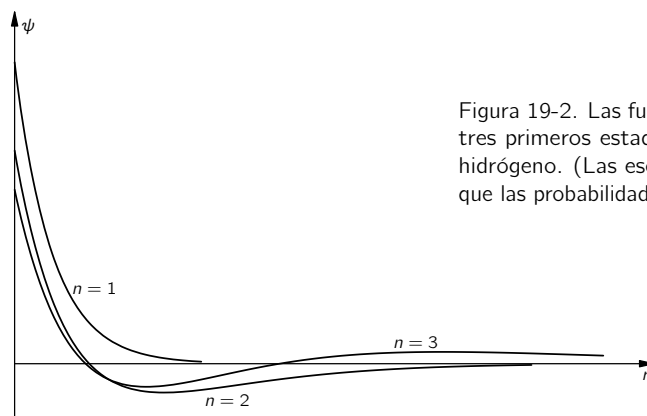


Figura 19-2. Las funciones de onda para los tres primeros estados $l = 0$ del átomo de hidrógeno. (Las escalas se eligen de modo que las probabilidades totales sean iguales.)

19-3 Estados con dependencia angular

En los estados descritos por $\psi_n(r)$ hemos encontrado que la amplitud de probabilidad de encontrar el electrón es esféricamente simétrica —dependiendo solamente de r , distancia al protón—. Tales estados tienen momentum angular orbital cero. Deberíamos ahora averiguar algo acerca de los estados que pueden tener algunas dependencias angulares.

Podríamos, si deseásemos, investigar solamente el problema estrictamente matemático de encontrar las funciones de r, θ y ϕ que satisfacen la ecuación diferencial (19.7) —con las condiciones físicas adicionales de que las únicas funciones aceptables son las que tienden a cero para r grandes—. Esto lo encontrarán así en muchos libros. Nosotros vamos a acortar un poco haciendo

uso del conocimiento que ya poseemos acerca de cómo dependen las amplitudes de los ángulos en el espacio.

El átomo de hidrógeno en cualquier estado determinado es una partícula con un cierto “espín” j —número cuántico del momentum angular total—. Parte de este espín proviene del espín intrínseco del electrón y parte del movimiento del electrón. Como cada una de estas componentes actúa independientemente (con una buena aproximación), ignoraremos de nuevo la parte del espín y consideraremos solamente el momentum “orbital”. Este movimiento orbital se comporta, sin embargo, del mismo modo que un espín. Por ejemplo, si el número cuántico orbital es l , la componente z del momentum angular puede ser $l, l-1, l-2, \dots, -l$. (Estamos, como de costumbre, midiendo en unidades de $\hbar$.) También sirven todas las matrices de rotación y las demás propiedades que calculamos (de ahora en adelante, ignoraremos *realmente* el espín del electrón; cuando hablemos de “momentum angular” entenderemos solamente la parte orbital).

Como el potencial V en el que se mueve el electrón depende solamente de r y no de θ o ϕ , el hamiltoniano es simétrico frente a cualquier rotación. De ahí se sigue que se conserva el momentum angular y todas sus componentes. (Esto es válido para un movimiento en *cualquier* “campo central” —que depende solamente de r — y no es una característica especial del potencial coulombiano e^2/r .)

Consideremos ahora algunos estados posibles del electrón; su estructura angular interna estará caracterizada por el número cuántico l . Dependiendo de la “orientación” del momentum angular total respecto al eje z , la componente z del momentum angular será m , una de las $2l+1$ posibilidades entre $+l$ y $-l$. Digamos $m=1$. ¿Con qué amplitud se encontrará el electrón en el eje z a una distancia r ? Cero. Un electrón en el eje z *no puede* tener ningún momentum angular orbital alrededor de ese eje. Muy bien, pero supongan que m es cero y, entonces, puede haber cierta amplitud distinta de cero de encontrar el electrón a cualquier distancia del protón. Llamaremos a esta amplitud $F_l(r)$. Es la amplitud de encontrar el electrón a la distancia r a lo largo de z positivo cuando el átomo está en el estado $|l, 0\rangle$, con lo cual entendemos espín orbital l y la componente z , con $m=0$.

Si conocemos $F_l(r)$ lo conocemos todo. Para cualquier estado $|l, m\rangle$, conocemos la amplitud $\psi_{l,m}(r)$ de encontrar al electrón en *cualquier* lugar del átomo. ¿Cómo? Observen. Supongan que tenemos el átomo en el estado $|l, m\rangle$, ¿cuál es la amplitud de encontrar al electrón al ángulo θ, ϕ y a la distancia r del origen? Hagan un nuevo eje z, z' digamos, en su ángulo (véase la figura 19-3), y pregunten: ¿cuál es la amplitud de que el electrón esté a la distancia r según el nuevo eje z' ? Sabemos que no se puede encontrar a lo largo del eje z' , a no ser que su componente z' del momentum angular, digamos m' , sea cero. Sin embargo, cuando m' es cero, la amplitud de encontrar el electrón según z' es $F_l(r)$. Así, el resultado es el producto de dos factores. El primero es la amplitud de que un átomo en el estado $|l, m\rangle$ según el eje z esté en el estado $|l, m'=0\rangle$, *respecto al eje z'* . Multipliquen su amplitud por $F_l(r)$ y tienen la amplitud $\psi_{l,m}(\mathbf{r})$ de encontrar al electrón en (r, θ, ϕ) respecto a los ejes originales.

Escribámoslo. Hemos calculado antes las matrices de transformación para rotaciones. Para ir del sistema x, y, z al x', y', z' de la figura 19-3, podemos en primer lugar rotar el eje z en un ángulo ϕ y luego en un ángulo θ alrededor del *nuevo* eje (y'). Esta rotación combinada es el producto

$$R_y(\theta)R_z(\phi).$$

La amplitud de encontrar el estado $l, m'=0$ después de la rotación es

$$\langle l, 0 | R_y(\theta)R_z(\phi) | l, m \rangle. \quad (19.31)$$

Nuestro resultado es, pues

$$\psi_{l,m}(\mathbf{r}) = \langle l, 0 | R_y(\theta)R_z(\phi) | l, m \rangle F_l(r). \quad (19.32)$$

El movimiento orbital puede tener solamente valores enteros de l (si el electrón se puede encontrar en cualquier lugar $r \neq 0$, hay cierta amplitud de tener $m=0$

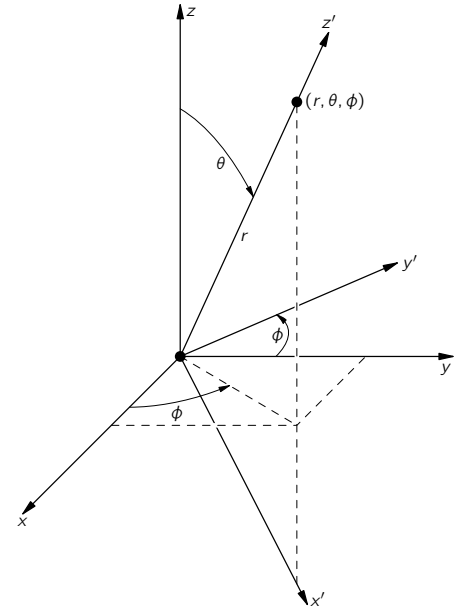


Figura 19-3. El punto (r, θ, ϕ) está sobre el eje z' de las coordenadas x', y', z' .

en esa dirección. Y los estados $m = 0$ existen solamente para espines enteros). Las matrices de rotación para $l = 1$ están en la tabla 17-2. Para l grandes pueden usar las fórmulas generales del capítulo 18. Las matrices $R_z(\phi)$ y $R_y(\theta)$ aparecen separadas, pero saben cómo combinarlas. Para el caso general, empezarán con el estado $|l, m\rangle$ y aplicarán $R_z(\phi)$ para obtener el nuevo estado $R_z(\phi)|l, m\rangle$ (que es justamente $e^{im\phi}|l, m\rangle$). Aplican, entonces, $R_y(\theta)$ a este estado y obtienen el estado $R_y(\theta)R_z(\phi)|l, m\rangle$. Al multiplicar por $\langle l, 0|$ nos da el elemento de matriz (19.31).

Los elementos de matriz de la operación de rotación son funciones algebraicas de θ y ϕ . Las funciones particulares que aparecen en (19.31) también entran en muchas clases de problemas que implican ondas en geometría esférica y por eso reciben un nombre especial. No todo el mundo usa la misma convención; pero una de las más comunes es

$$\langle l, 0 | R_y(\theta)R_z(\phi) | l, m \rangle \equiv a Y_{l,m}(\theta, \phi). \quad (19.33)$$

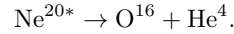
Las funciones $Y_{l,m}(\theta, \phi)$ se llaman *armónicos esféricos* y a es solamente un factor numérico que depende de la definición escogida para $Y_{l,m}$. Usualmente se define,

$$a = \sqrt{\frac{4\pi}{2l+1}}. \quad (19.34)$$

Con esta notación, las funciones de onda del hidrógeno se pueden escribir

$$\psi_{l,m}(\mathbf{r}) = a Y_{l,m}(\theta, \phi) F_l(r). \quad (19.35)$$

Las funciones angulares $Y_{l,m}(\theta, \phi)$ no solamente son importantes en muchos problemas cuánticos, sino también en muchas áreas de la física clásica en que aparece el operador ∇^2 , tal como en el electromagnetismo. Como otro ejemplo de su uso en mecánica cuántica, consideren la desintegración de un estado excitado de Ne^{20} (como se discutió en el capítulo anterior), que se desintegra emitiendo una partícula α convirtiéndose en O^{16} :



Supongan que el estado excitado tiene algún espín l (necesariamente entero) y que la componente z del momentum angular es m . Podríamos preguntar lo siguiente: dados l y m , ¿cuál es la amplitud de que encontremos la partícula α saliendo en una dirección que forma el ángulo θ respecto al eje z y el ángulo ϕ respecto al plano xz , como muestra la figura 19-4?

Para resolver este problema, hacemos primero la siguiente observación. Una desintegración en la que la partícula α marcha derecho a lo largo del eje z positivo debe proceder de un estado con $m = 0$. Esto es así porque tanto O^{16} como la partícula α tienen espín cero y porque sus movimientos no pueden tener ningún momentum angular respecto al eje z . Llamemos a esta amplitud a (por unidad de ángulo sólido). Entonces, para encontrar la amplitud de una desintegración en el ángulo arbitrario de la figura 19-4, todo lo que necesitamos saber es a qué amplitud el estado inicial dado tiene momentum cero respecto a la dirección de desintegración. La amplitud para la desintegración en θ y ϕ es entonces a por la amplitud del estado $|l, m\rangle$, respecto al eje z que esté en el estado $|l, 0\rangle$ respecto a z' —dirección de la desintegración—. Esta última amplitud es lo que exactamente hemos escrito en (19.31). La probabilidad de ver la partícula α en θ, ϕ es

$$P(\theta, \phi) = a^2 |\langle l, 0 | R_y(\theta)R_z(\phi) | l, m \rangle|^2.$$

Como un ejemplo, consideren un estado inicial con $l = 1$ y diversos valores de m . La tabla 17-2 nos da las amplitudes necesarias. Son:

$$\begin{aligned} \langle 1, 0 | R_y(\theta)R_z(\phi) | 1, +1 \rangle &= -\frac{1}{\sqrt{2}} \sin \theta e^{i\phi}, \\ \langle 1, 0 | R_y(\theta)R_z(\phi) | 1, 0 \rangle &= \cos \theta, \\ \langle 1, 0 | R_y(\theta)R_z(\phi) | 1, -1 \rangle &= \frac{1}{\sqrt{2}} \sin \theta e^{-i\phi}. \end{aligned} \quad (19.36)$$

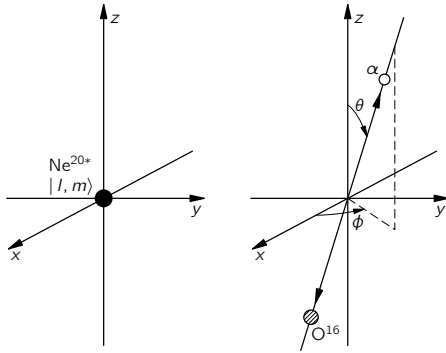


Figura 19-4. La descomposición de un estado excitado de Ne^{20} .

Éstas son las tres amplitudes posibles de distribución angular —que dependen del valor de m del núcleo inicial—.

Amplitudes como las de (19.36) aparecen con frecuencia y son tan importantes que se les han dado *diversos* nombres. Si la amplitud de distribución angular es proporcional a cualquiera de las tres funciones o a cualquier combinación lineal de ellas, decimos: “el sistema tiene un momentum angular orbital de uno”. O podemos decir: “el Ne^{20*} emite una partícula α en una onda p ”. O decimos: “la partícula α es emitida en un estado $l = 1$ ”. Como hay tantos modos de decir lo mismo, un diccionario resulta útil. Si quieren entender lo que otros físicos estén hablando, tienen necesariamente que aprenderse de memoria este lenguaje. En la tabla 19-1 damos un diccionario del momentum angular orbital.

Si el momentum angular orbital es cero, no hay cambio cuando se rota el sistema de coordenadas y no hay variación con el ángulo —la “dependencia” angular es como una constante 1, digamos—. A esto se le llama también un “estado s ” y hay solamente uno —en lo que respecta a la dependencia angular—. Si el momentum angular orbital es 1, la amplitud de la variación angular puede ser cualquiera de las tres funciones dadas —dependiendo del valor de m — o puede ser una combinación lineal. Se les llama “estados p ” y hay tres. Si el momentum angular orbital es 2, hay las cinco funciones que se muestran. Cualquier combinación lineal se llama una amplitud “ $l = 2$ ” o una amplitud “de onda d ”. Ahora pueden adivinar la letra siguiente —¿qué debe venir después de s , p , d ?—. Bueno, ¡naturalmente, f , g , h y las restantes del alfabeto! Las letras no significan nada. (Una vez significaron algo —significaron líneas “nítidas”, líneas “principales”, líneas “difusas” y líneas “fundamentales” de los espectros ópticos de los átomos—. Pero esto era en los tiempos en que la gente no sabía de dónde procedían las líneas. Después de f no había nombres especiales, por lo que ahora nosotros continuamos simplemente con g , h , etcétera.)

Las funciones angulares de la tabla tienen diferentes nombres —y se definen algunas veces con convenciones levemente diferentes acerca de los factores numéricos que aparecen delante—. Algunas veces se les llama “armónicos esféricos” y se escriben como $Y_{l,m}(\theta, \phi)$. Otras veces se escriben $P_l^m(\cos \theta)e^{im\phi}$, y si $m = 0$ simplemente como $P_l(\cos \theta)$. Las funciones $P_l(\cos \theta)$ se llaman “polinomios de Legendre” en $\cos \theta$, y las funciones $P_l^m(\cos \theta)$ “funciones de Legendre asociadas”. En muchos libros encontrarán tablas de estas funciones.

Observen, entre paréntesis, que todas las funciones para un l dado tienen la propiedad de tener la misma paridad —para valores impares de l cambian el signo frente a una inversión y para l pares no cambian—. Así, podemos escribir que la paridad de un estado de momentum angular orbital l es $(-1)^l$.

Como hemos visto, estas distribuciones angulares se pueden referir a la desintegración nuclear o a cualquier otro proceso, o a la distribución de amplitud de encontrar un electrón en algún lugar en el átomo de hidrógeno. Por ejemplo, si un electrón está en un estado p ($l = 1$), la amplitud de encontrarlo puede depender del ángulo de muchas maneras posibles —pero todas son combinaciones lineales de las tres funciones para $l = 1$ de la tabla 19-1—. Tomemos el caso de $\cos \theta$. Es interesante. Significa que la amplitud es positiva, digamos, en la parte superior ($\theta < \pi/2$), es negativa en la parte inferior ($\theta > \pi/2$) y es cero cuando θ es 90° . Elevando al cuadrado esta amplitud vemos que la probabilidad de encontrar el electrón varía con θ , como muestra la figura 19-5 —y es independiente de ϕ —. Esta distribución angular es responsable del hecho de que en la ligadura molecular la atracción de un electrón en un estado $l = 1$ hacia otro átomo depende de la dirección —es el origen de las valencias dirigidas de la atracción química—.

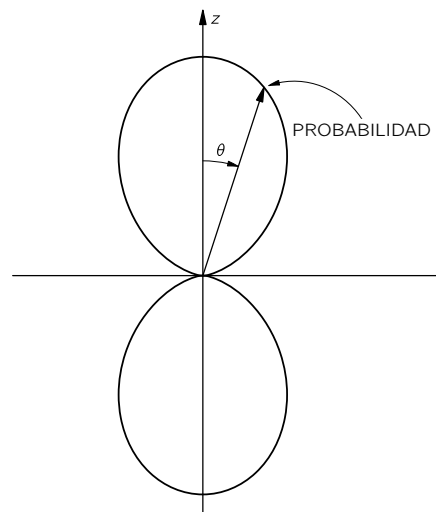


Figura 19-5. Gráfica de $\cos^2 \theta$ en coordenadas polares, es la probabilidad relativa de encontrar un electrón en varios ángulos desde el eje z (dado un r) en un estado atómico, con $l = 1$ y $m = 0$.

19-4 Solución general para el hidrógeno

En la ecuación (19.35) hemos escrito las funciones de onda del átomo de hidrógeno en la forma:

$$\psi_{l,m}(\mathbf{r}) = a Y_{l,m}(\theta, \phi) F_l(r). \quad (19.37)$$

Tabla 19-1

Diccionario del momentum angular orbital
($l = j = \text{un número entero}$)

Momentum angular orbital, l	Componente z, m	Dependencia angular de amplitudes	Nombre	Número de estados	Paridad orbital
0	0	1	s	1	+
1	$\begin{Bmatrix} +1 \\ 0 \\ -1 \end{Bmatrix}$	$\begin{Bmatrix} -\frac{1}{\sqrt{2}} \text{sen } \theta e^{i\phi} \\ \cos \theta \\ \frac{1}{\sqrt{2}} \text{sen } \theta e^{-i\phi} \end{Bmatrix}$	p	3	−
2	$\begin{Bmatrix} +2 \\ +1 \\ 0 \\ -1 \\ -2 \end{Bmatrix}$	$\begin{Bmatrix} \frac{\sqrt{6}}{4} \text{sen}^2 \theta e^{2i\phi} \\ \frac{\sqrt{6}}{2} \text{sen } \theta \cos \theta e^{i\phi} \\ \frac{1}{2} (3 \cos^2 \theta - 1) \\ -\frac{\sqrt{6}}{2} \text{sen } \theta \cos \theta e^{-i\phi} \\ \frac{\sqrt{6}}{4} \text{sen}^2 \theta e^{-2i\phi} \end{Bmatrix}$	d	5	+
3	$\left. \begin{Bmatrix} \\ \\ \\ \end{Bmatrix} \right\}$	$\left. \begin{Bmatrix} \langle l, 0 R_y(\theta) R_z(\phi) l, m \rangle \\ = Y_{l,m}(\theta, \phi) \\ = P_l^m(\cos \theta) e^{im\phi} \end{Bmatrix} \right\}$	f	$2l + 1$	$(-1)^l$
4			g		
5			h		
$\vdots$			$\vdots$		

Estas funciones de onda deben ser soluciones de la ecuación diferencial (19.7). Veamos lo que esto significa. Introduzcan (19.37) en (19.7); obtienen:

$$\begin{aligned} \frac{Y_{l,m}}{r} \frac{\partial^2}{\partial r^2} (r F_l) + \frac{F_l}{r^2 \text{sen } \theta} \frac{\partial}{\partial \theta} \left(\text{sen } \theta \frac{\partial Y_{l,m}}{\partial \theta} \right) + \frac{F_l}{r^2 \text{sen}^2 \theta} \frac{\partial^2 Y_{l,m}}{\partial \phi^2} \\ = -\frac{2m}{\hbar^2} \left(E + \frac{e^2}{r} \right) Y_{l,m} F_l. \quad (19.38) \end{aligned}$$

Multipliquen ahora todo por r^2/F_l y reorganicen los términos. El resultado es:

$$\begin{aligned} \frac{1}{\text{sen } \theta} \frac{\partial}{\partial \theta} \left(\text{sen } \theta \frac{\partial Y_{l,m}}{\partial \theta} \right) + \frac{1}{\text{sen}^2 \theta} \frac{\partial^2 Y_{l,m}}{\partial \phi^2} \\ = - \left[\frac{r^2}{F_l} \left\{ \frac{1}{r} \frac{d^2}{dr^2} (r F_l) + \frac{2m}{\hbar^2} \left(E + \frac{e^2}{r} \right) F_l \right\} \right] Y_{l,m}. \quad (19.39) \end{aligned}$$

El primer miembro de esta ecuación depende de θ y de ϕ , *pero no de r* . El primer miembro no varía, cualquiera sea el valor que elijamos para r . *Esto debe ser también válido para el segundo miembro*. Aunque la cantidad entre corchetes tenga r en todas partes, la cantidad total no puede depender de r porque de otro modo no tendríamos una ecuación válida para cualquier r . Como pueden ver, el corchete tampoco depende de θ o ϕ . Debe ser algo constante. Su valor puede depender del valor l del estado que estamos estudiando, puesto que la función F_l debe ser la apropiada para ese estado; la llamaremos constante K_l . La ecuación (19.39) es, por consiguiente, equivalente a *dos* ecuaciones:

$$\frac{1}{\sin \theta} \frac{\partial}{\partial \theta} \left(\sin \theta \frac{\partial Y_{l,m}}{\partial \theta} \right) + \frac{1}{\sin^2 \theta} \frac{\partial^2 Y_{l,m}}{\partial \phi^2} = -K_l Y_{l,m}, \quad (19.40)$$

$$\frac{1}{r} \frac{d^2}{dr^2} (r F_l) + \frac{2m}{\hbar^2} \left(E + \frac{e^2}{r} \right) F_l = K_l \frac{F_l}{r^2}. \quad (19.41)$$

Observen lo que hemos hecho. Para cualquier estado descrito por l y m , conocemos las funciones $Y_{l,m}$; podemos usar la ecuación (19.40) para determinar la constante K_l . Introduciendo K_l en (19.41) tenemos una ecuación diferencial para la función $F_l(r)$. Si podemos resolver esta ecuación en $F_l(r)$, tenemos todo lo necesario para que (19.37) nos dé $\psi(\mathbf{r})$.

¿Qué es K_l ? En primer lugar, observen que debe ser la misma para toda m (la cual acompaña a un l particular), por lo que podemos escoger la m que deseemos para $Y_{l,m}$ y la introducimos en (19.40) para hallar K_l . Quizá la más fácil de usar es $Y_{l,l}$. De la ecuación (18.24),

$$R_z(\phi) |l, l\rangle = e^{il\phi} |l, l\rangle. \quad (19.42)$$

El elemento de matriz de $R_y(\theta)$ es también bastante simple:

$$\langle l, 0 | R_y(\theta) | l, l \rangle = b (\sin \theta)^l, \quad (19.43)$$

donde b es cierto número.[†] Combinando las dos, obtenemos:

$$Y_{l,l} \propto e^{il\phi} \sin^l \theta. \quad (19.44)$$

Introduciendo esta función en (19.40) nos da:

$$K_l = l(l+1). \quad (19.45)$$

Ahora que hemos determinado K_l , la ecuación (19.41) nos cuenta cómo es la función radial $F_l(r)$. Es, naturalmente, la ecuación de Schrödinger exacta con la parte angular remplazada por su equivalente $K_l F_l / r^2$. Volvamos a escribir (19.41) en la forma que teníamos en la ecuación (19.8), como sigue:

$$\frac{1}{r} \frac{d^2}{dr^2} (r F_l) = -\frac{2m}{\hbar^2} \left\{ E + \frac{e^2}{r} - \frac{l(l+1)\hbar^2}{2mr^2} \right\} F_l. \quad (19.46)$$

Un término misterioso se ha añadido a la energía potencial. Aunque obtuvimos este término mediante algunas maquinaciones matemáticas, tiene un origen físico sencillo. Podemos darles una idea de dónde proviene en términos de un razonamiento semiclásico. Después quizá no lo encontrarán tan misterioso.

Tomen una partícula clásica que se mueve entorno de algún centro de fuerza. La energía total se conserva y es la suma de las energías potencial y cinética.

$$U = V(r) + \frac{1}{2}mv^2 = \text{constante}.$$

En general, v se puede descomponer en una componente radial v_r y una componente tangencial $r\dot{\theta}$; así

$$v^2 = v_r^2 + (r\dot{\theta})^2.$$

Pero el momentum angular $mr^2\dot{\theta}$ también se conserva; digamos que es igual a L . Podemos entonces escribir

$$mr^2\dot{\theta} = L, \quad \text{o} \quad r\dot{\theta} = \frac{L}{mr},$$

y la energía es

[†] Pueden demostrar con algún trabajo que esto proviene de la ecuación (18.35), pero también es fácil calcularlo a partir de los primeros principios siguiendo las ideas de la sección 18-4. Un estado $|l, l\rangle$ se puede formar de $2l$ partículas de espín un medio, todas con espín hacia arriba; mientras que el estado $|l, 0\rangle$ tendría l hacia arriba y l hacia abajo. Frente a una rotación, la amplitud de que un espín hacia arriba permanezca hacia arriba es $\cos \theta/2$ y de que un espín hacia arriba cambie hacia abajo es $-\sin \theta/2$. Estamos preguntando por la amplitud de que l espines hacia arriba permanezcan así, mientras que los otros l espines hacia arriba cambien hacia abajo. La amplitud de esto es $(\cos \theta/2 \sin \theta/2)^l$ que es proporcional a $\sin^l \theta$.

$$U = \frac{1}{2}mv_r^2 + V(r) + \frac{L^2}{2mr^2}.$$

Si no hubiese ningún momentum angular, sólo tendríamos los dos primeros términos. Añadir el momentum angular L afecta a la energía exactamente lo mismo que afectaría añadir un término $L^2/2mr^2$ a la energía potencial. Pero éste es casi exactamente el término adicional en (19.46). La sola diferencia es que $l(l+1)\hbar^2$ aparece como momentum angular en lugar de $l^2\hbar^2$ como podríamos esperar. Pero vimos antes (por ejemplo, véase la sección 34-7 del volumen II)[†] que ésta es exactamente la sustitución requerida para hacer que un razonamiento casi clásico concuerde con un cálculo cuántico correcto. Podemos, pues, entender el nuevo término como un “pseudopotencial” que da el término de “fuerza centrífuga” que aparece en las ecuaciones de movimiento radial para un sistema que rota. (Revisen la discusión de “pseudofuerzas” en la sección 12-5 del volumen I.)

Ahora estamos preparados para resolver la ecuación (19.46) en $F_l(r)$. Se parece mucho a (19.8), por lo que servirá la misma técnica. Todo marcha como antes hasta que llegan a la ecuación (19.19) que tendrá el término adicional

$$-l(l+1) \sum_{k=1}^{\infty} a_k \rho^{k-2}. \quad (19.47)$$

Este término también se puede escribir en la forma

$$-l(l+1) \left\{ \frac{a_1}{\rho} + \sum_{k=1}^{\infty} a_{k+1} \rho^{k-1} \right\}. \quad (19.48)$$

y la energía es

$$\sum_{k=1}^{\infty} [\{k(k+1) - l(l+1)\}a_{k+1} - 2(\alpha k - 1)a_k] \rho^{k-1} - \frac{l(l+1)a_1}{\rho} = 0. \quad (19.49)$$

Hay sólo un término en ρ^{-1} , por lo que debe ser cero. El coeficiente a_1 debe ser cero (a no ser que $l = 0$ y tenemos la solución anterior). Cada uno de los otros términos se hace cero al obtener el corchete para toda k . Esta condición reemplaza la ecuación (19.22) por

$$a_{k+1} = \frac{2(\alpha k - 1)}{k(k+1) - l(l+1)} a_k. \quad (19.50)$$

Éste es el único cambio significativo respecto al caso esféricamente simétrico.

Como antes, la serie debe ser finita si queremos soluciones que representen electrones ligados. La serie terminará para $k = n$ si $\alpha n = 1$. Obtenemos de nuevo la misma condición para α , debe ser igual a $1/n$, donde n es un entero positivo. Sin embargo, la ecuación (19.50) también impone una nueva restricción. El índice k no puede ser igual a l , porque el denominador se hace cero y a_{l+1} sería infinito. Esto es, como $a_1 = 0$, la ecuación (19.50) implica que todas las siguientes a_k son cero hasta que llegamos a a_{l+1} , que puede ser distinto de cero. Esto significa que k debe empezar en $l+1$ y terminar en n .

Nuestro resultado final es que para cualquier l hay muchas soluciones posibles que podemos llamar $F_{n,l}$ con $n \geq l+1$. Cada solución tiene la energía

$$E_n = -\frac{me^4}{2\hbar^2} \left(\frac{1}{n^2} \right). \quad (19.51)$$

La función de onda para el estado con esta energía y números cuánticos angulares l y m es

$$\psi_{n,l,m} = a Y_{l,m}(\theta, \phi) F_{n,l}(\rho), \quad (19.52)$$

con

[†] Véase el apéndice de este volumen.

$$\rho F_{n,l}(\rho) = e^{-\alpha\rho} \sum_{k=l+1}^n a_k \rho^k. \quad (19.53)$$

Los coeficientes a_k se obtienen de (19.50). Tenemos, finalmente, una descripción completa de los estados de un átomo de hidrógeno.

19-5 Funciones de onda del hidrógeno

Revisemos lo que hemos descubierto. Los estados que satisfacen la ecuación de Schrödinger para un electrón en un campo coulombiano están caracterizados por tres números cuánticos n , l y m , todos enteros. La distribución angular de la amplitud del electrón puede tener solamente ciertas formas que llamamos $Y_{l,m}$. Están clasificadas con l , *número cuántico del momentum angular total*, y m , *número cuántico “magnético”*, que puede ir desde $-l$ hasta $+l$. Para cada configuración angular, hay varias distribuciones radiales posibles $F_{n,l}(r)$ de la amplitud del electrón; se designa por el *número cuántico principal* n —que va desde $l+1$ hasta ∞ —. La energía del estado depende solamente de n y se incrementa cuando n aumenta.

El estado de energía más baja, o estado fundamental, es un estado s . Tiene $l=0$, $n=1$ y $m=0$. Es un estado “no degenerado” —hay solamente uno con esta energía y su función de onda es esféricamente simétrica—. La amplitud de encontrar el electrón tiene un máximo en el centro y cae monótonamente cuando aumenta la distancia al centro. Podemos imaginar la amplitud del electrón como una burbuja, según muestra la figura 19-6(a).

Hay otros estados s con energías más altas, para $n=2, 3, 4, \dots$. Para cada energía hay sólo una versión ($m=0$) y todos son esféricamente simétricos. Estos estados tienen amplitudes que cambian de signo una o más veces al aumentar r . Hay $n-1$ superficies nodales esféricas —los lugares donde ψ pasa por cero—. El estado $2s$ ($l=0$, $n=2$), por ejemplo, se parecerá a lo dibujado en la figura 19-6(b) (las áreas oscuras indican regiones donde la amplitud es grande y los signos más y menos indican las fases relativas de la amplitud). Los niveles de energía de los estados s están en la primera columna de la figura 19-7.

Luego vienen los estados p —con $l=1$ —. Para cada n , que debe ser 2 o mayor, hay tres estados con igual energía, cada uno de ellos para $m=+1$, $m=0$ y $m=-1$. La figura 19-7 muestra estos niveles de energía. Las dependencias angulares de estos estados están en la tabla 19-1. Por ejemplo, para $m=0$, si la amplitud es positiva para θ casi cero, será negativa para θ cerca de 180° . Hay un plano nodal coincidente con el plano xy . Para $n>2$ hay también nodos esféricos. La amplitud $n=2$, $m=0$ se ha representado en la figura 19-6(c) y la función de onda $n=3$, $m=0$ en la figura 19-6(d).

Pueden pensar que como m representa un tipo de “orientación” en el espacio, debería haber distribuciones similares con los picos de amplitud según el eje x o según el eje y . ¿Son éstos, quizá, los estados $m=+1$ y $m=-1$? No. Pero como tenemos tres estados de igual energía, cualquier combinación lineal de los mismos también será estado estacionario con la misma energía. Resulta que el estado “ x ” —que corresponde al estado “ z ”, o estado $m=0$, de la figura 19-6(c)— es una combinación lineal de los estados $m=+1$ y $m=-1$. El estado correspondiente “ y ” es otra combinación. Específicamente queremos decir que

$$\begin{aligned} \text{“}z\text{”} &= |1, 0\rangle, \\ \text{“}x\text{”} &= -\frac{|1, +1\rangle - |1, -1\rangle}{\sqrt{2}}, \\ \text{“}y\text{”} &= -\frac{|1, +1\rangle + |1, -1\rangle}{i\sqrt{2}}. \end{aligned}$$

Todos estos estados tienen se ven igual cuando se refieren a sus ejes respectivos.

Los estados d ($l=2$) tienen cinco valores posibles de m para cada energía, siendo la energía más baja para $n=3$. Los niveles son como en la figura 19-7.

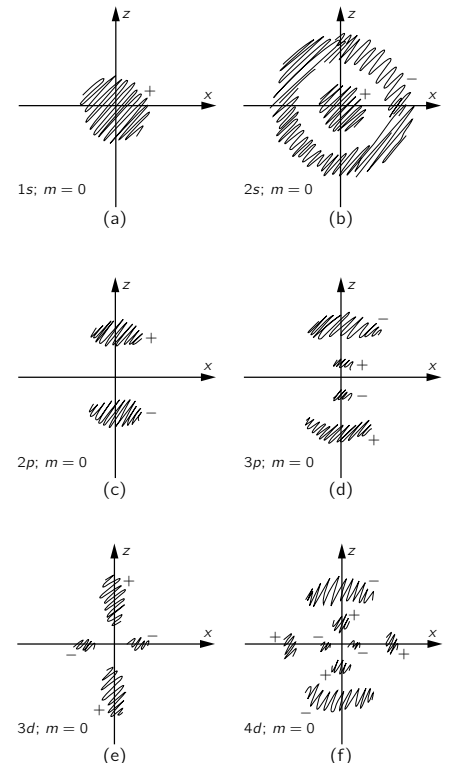


Figura 19-6. Esbozos que muestran la naturaleza general de algunas de las funciones de onda del hidrógeno. Las regiones sombreadas muestran dónde las amplitudes son grandes. Los signos más y menos muestran el signo relativo de la amplitud en cada región.

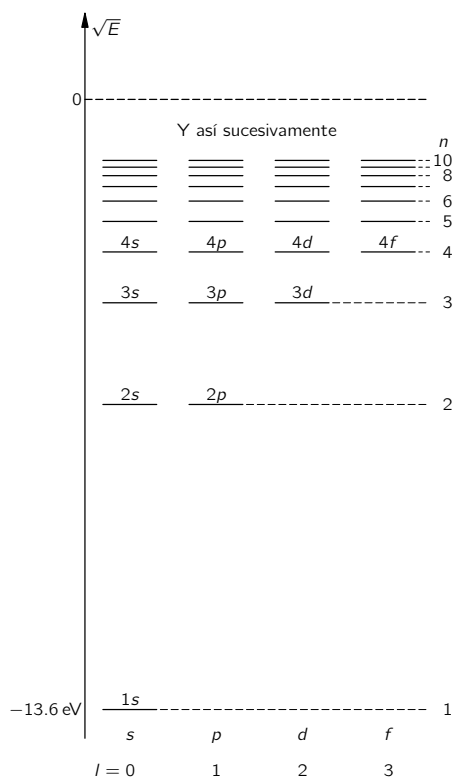


Figura 19-7. Diagrama de los niveles de energía del hidrógeno.

Las dependencias angulares son más complicadas. Por ejemplo, los estados $m = 0$ tienen dos nodos cónicos, por lo que la función de onda invierte la fase de $+$, a $-$, a $+$ cuando da la vuelta desde el polo norte al polo sur. La forma aproximada de la amplitud está representada en (e) y (f) de la figura 19-6 para los estados $m = 0$ con $n = 3$ y $n = 4$. De nuevo, las n grandes tienen nodos esféricos.

No trataremos de describir ninguno más de los estados posibles. Encontrarán las funciones de onda del hidrógeno detalladamente descritas en muchos libros. Dos buenas referencias son L. Pauling y E. B. Wilson, *Introduction to Quantum Mechanics*, McGraw-Hill (1935) y R. B. Leighton, *Principles of Modern Physics*, McGraw-Hill (1959). Encontrarán en ellos gráficas de algunas de las funciones y representaciones gráficas de muchos estados.

Queríamos mencionar una característica particular de las funciones de onda para l más grandes; para $l > 0$ las amplitudes son cero en el centro. Esto no es sorprendente, ya que es difícil que un electrón tenga momentum angular cuando su distancia radial es muy pequeña. Por esta razón, cuanto más grande sea l tanto más “se apartan” las amplitudes del centro. Si consideran el modo en que las funciones radiales $F_{n,l}(r)$ varían con r pequeñas, encuentran a partir de (19.53) que

$$F_{n,l}(r) \approx r^l.$$

Tal dependencia de r significa que para l grandes tienen que apartarse más de $r = 0$ antes de que obtengan una amplitud apreciable. Incidentalmente, este comportamiento está determinado por el término de fuerza centrífuga en la ecuación radial, por lo cual lo mismo servirá para cualquier potencial que varíe más lentamente que $1/r^2$ para r pequeñas —lo que hacen la mayoría de los potenciales atómicos—.

19-6 La tabla periódica

Ahora nos gustaría aplicar la teoría del átomo de hidrógeno de un modo aproximado para alcanzar a comprender algo de la tabla periódica de los elementos que usan los químicos. Para un elemento con número atómico Z , hay Z electrones mantenidos juntos por la atracción eléctrica del núcleo, pero con una repulsión mutua de los electrones. Para obtener una solución exacta tendríamos que resolver la ecuación de Schrödinger para Z electrones en un campo coulombiano. Para el helio la ecuación es:

$$-\frac{\hbar}{i} \frac{\partial \psi}{\partial t} = -\frac{\hbar^2}{2m} (\nabla_1^2 \psi + \nabla_2^2 \psi) + \left(-\frac{2e^2}{r_1} - \frac{2e^2}{r_2} + \frac{e^2}{r_{12}} \right) \psi,$$

donde ∇_1^2 es un laplaciano que opera sobre $\mathbf{r}_1$, coordenada de un electrón; ∇_2^2 opera sobre $\mathbf{r}_2$, y $r_{12} = |\mathbf{r}_1 - \mathbf{r}_2|$ (de nuevo estamos despreciando el espín de los electrones). Para encontrar los estados estacionarios y los niveles de energía tendríamos que encontrar soluciones de la forma:

$$\psi = f(\mathbf{r}_1, \mathbf{r}_2) e^{-(i/\hbar)Et}.$$

La dependencia geométrica está incluida en f , que es una función de seis variables —las posiciones simultáneas de dos electrones—. Nadie ha encontrado una solución analítica, aunque se han obtenido soluciones para los estados de energía más bajos, mediante métodos numéricos.

Con 3, 4 o 5 electrones es inútil tratar de obtener soluciones exactas, y es mucho decir que la mecánica cuántica ha dado un conocimiento preciso de la tabla periódica. Sin embargo, es posible aunque sea con una aproximación más bien burda —y algunos ajustes— entender, al menos cualitativamente, muchas propiedades químicas que muestra la tabla periódica.

Las propiedades químicas de los átomos están determinadas primariamente por sus estados de energía más bajos. Podemos usar la siguiente teoría aproximada para encontrar estos estados y sus energías. Primero, despreciamos el espín del electrón, *excepto* que adoptamos el principio de exclusión y decimos que

cualquier estado electrónico particular puede ser ocupado solamente por un electrón. Esto significa que cualquier configuración orbital particular puede tener *dos* electrones —uno con espín hacia arriba y otro con espín hacia abajo—. Luego descartamos los *detalles* de las interacciones entre los electrones en nuestra primera aproximación y decimos que cada electrón se mueve en un *campo central* que es el campo combinado del núcleo y de los otros electrones. Para el neón, que tiene 10 electrones, decimos que un electrón ve un potencial promedio debido al núcleo más los otros nueve electrones. Imaginamos, luego, que en la ecuación de Schrödinger para cada electrón ponemos un $V(r)$ que es un campo $1/r$ modificado por una densidad de carga esféricamente simétrica que proviene de los otros electrones.

En este modelo cada electrón actúa como una partícula independiente. La dependencia angular de su función de onda será exactamente igual a las que obtuvimos para el átomo de hidrógeno. Habrá estados s , estados p , etc., y tendrán los diversos valores posibles de m . Como $V(r)$ ya no varía como $1/r$, la parte radial de las funciones de onda será algo diferente, pero será cualitativamente lo mismo, por lo que tendremos los mismos números cuánticos radiales n . Las energías de los estados también serán un poco diferentes.

H

Con estas ideas, veamos lo que obtenemos. El estado fundamental del hidrógeno tiene $l = m = 0$ y $n = 1$; decimos que la configuración del electrón es $1s$. La energía es -13.6 eV. Esto significa que necesitamos 13.6 electronvoltios para sacar el electrón del átomo. La llamamos “energía de ionización”, W_I . Una energía grande de ionización significa que es más difícil sacar el electrón y, en general, que el material es menos activo químicamente.

He

Tomemos ahora el helio. Ambos electrones pueden estar en el mismo estado inferior (uno con espín hacia arriba y otro con espín hacia abajo). Cuando se encuentra en este estado inferior, el electrón se mueve en un potencial que es para r pequeñas como un campo coulombiano para $z = 2$ y para r grandes como un campo coulombiano para $z = 1$. El resultado es un estado $1s$ “hidrogenoide” con una energía un poco más baja. Ambos electrones ocupan estados $1s$ idénticos ($l = 0, m = 0$). La energía de ionización observada (para quitar *un* electrón) es 24.6 electronvoltios. Como la “capa” $1s$ está ahora completa —permitimos solamente dos electrones— no hay prácticamente tendencia para que un electrón sea atraído por otro átomo. El helio es inerte químicamente.

Li

El núcleo de litio tiene una carga de 3. Los estados electrónicos serán de nuevo parecidos a los del hidrógeno y los tres electrones ocuparán los tres niveles de energía más bajos. Dos irán a estados $1s$ y el tercero irá a un estado $n = 2$. Pero ¿con $l = 0$ o con $l = 1$? En el hidrógeno estos estados tienen la misma energía, pero en otros átomos no, por la siguiente razón. Recuerden que un estado $2s$ tiene cierta amplitud de estar cerca del núcleo mientras que el estado $2p$ no. Esto significa que un electrón $2s$ sentirá algo de la triple carga eléctrica del núcleo de Li, pero que un electrón $2p$ estará fuera donde el campo se parece al campo coulombiano de una sola carga. La atracción adicional disminuye la energía del estado $2s$ relativa al estado $2p$. Los niveles de energía serán aproximadamente como los que muestra la figura 19-8 —que deberían comparar con el diagrama correspondiente del hidrógeno en la figura 19-7—. Por lo tanto, el átomo de litio tendrá dos electrones en los estados $1s$ y uno en un $2s$. Como el electrón $2s$ tiene una energía más alta que un electrón $1s$ es relativamente fácil de sacar. La energía de ionización del litio es solamente 5.4 electronvoltios y es bastante activo químicamente.

Así pueden ver los diagramas que se van formando; hemos dado en la tabla 19-2 una lista de los 36 primeros elementos, con los estados ocupados por los

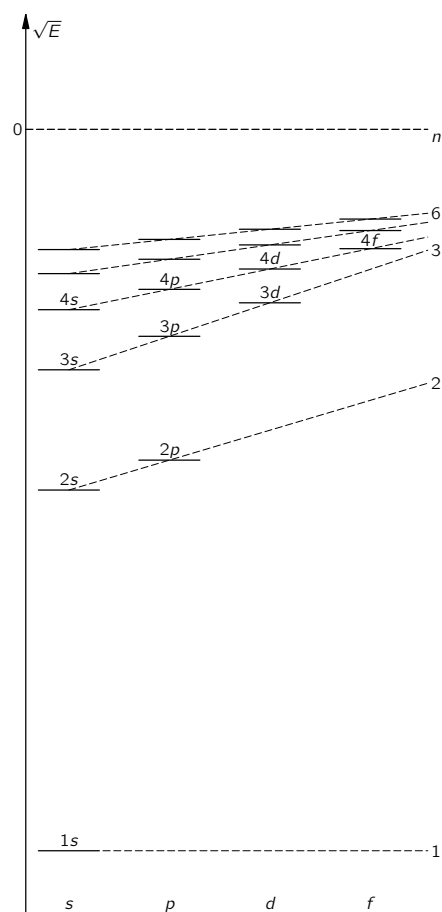


Figura 19-8. Diagrama esquemático del nivel de energía de un electrón atómico con otros electrones presentes. (La escala no es la misma que la de la figura 19-7.)

Tabla 19-2
Las configuraciones electrónicas de los primeros 36 elementos

Z	Elemento	$W_I(\text{eV})$	Electrón de configuración									
			1s	2s	2p	3s	3p	3d	4s	4p	4d	4f
1	H <i>hidrógeno</i>	13.6	1									
2	He <i>helio</i>	24.6	2									
3	Li <i>litio</i>	5.4		1								
4	Be <i>berilio</i>	9.3		2								
5	B <i>boro</i>	8.3		2	1							
6	C <i>carbón</i>	11.3	RELLENO (2)	2	2				Número de electrones en cada estado			
7	N <i>nitrógeno</i>	14.5		2	3							
8	O <i>oxígeno</i>	13.6		2	4							
9	F <i>flúor</i>	17.4		2	5							
10	Ne <i>neón</i>	21.6		2	6							
11	Na <i>sodio</i>	5.1				1						
12	Mg <i>magnesio</i>	7.6				2						
13	Al <i>aluminio</i>	6.0				2	1					
14	Si <i>silicio</i>	8.1	—RELLENO—			2	2					
15	P <i>fósforo</i>	10.5				2	3					
16	S <i>azufre</i>	10.4	(2)		(8)	2	4					
17	Cl <i>cloro</i>	13.0				2	5					
18	Ar <i>argón</i>	15.8				2	6					
19	K <i>potasio</i>	4.3							1			
20	Ca <i>calcio</i>	6.1							2			
21	Sc <i>escandio</i>	6.5						1	2			
22	Ti <i>titanio</i>	6.8						2	2			
23	V <i>vanadio</i>	6.7	—RELLENO—					3	2			
24	Cr <i>cromo</i>	6.8						5	1			
25	Mn <i>manganeso</i>	7.4	(2)		(8)	(8)		5	2			
26	Fe <i>hierro</i>	7.9						6	2			
27	Co <i>cobalto</i>	7.9						7	2			
28	Ni <i>níquel</i>	7.6						8	2			
29	Cu <i>cobre</i>	7.7						10	1			
30	Zn <i>zinc</i>	9.4						10	2			
31	Ga <i>galio</i>	6.0							2	1		
32	Ge <i>germanio</i>	7.9	—RELLENO—						2	2		
33	As <i>arsénico</i>	9.8							2	3		
34	Se <i>selenio</i>	9.7	(2)		(8)	(18)			2	4		
35	Br <i>bromo</i>	11.8							2	5		
36	Kr <i>criptón</i>	14.0							2	6		

electrones en el estado fundamental de cada átomo. La tabla da la energía de ionización para el electrón ligado más débilmente y el número de electrones que ocupan cada “capa” —con lo cual indicamos estados con el mismo n —. Como los diferentes estados l tienen energías diferentes, cada valor l corresponde a una subcapa con $2(2l + 1)$ estados posibles (de m y espín electrónico diferentes). Todos éstos tienen la misma energía; exceptuando algunos efectos muy pequeños que no estamos tomando en cuenta.

Be

El berilio es como el litio excepto que tiene dos electrones en el estado $2s$ y también, dos en la capa completa $1s$.

Be hasta Ne

El boro tiene cinco electrones. El quinto debe ir a un estado $2p$. Hay $2 \times 3 = 6$ estados $2p$ diferentes, por lo que podemos seguir añadiendo electrones hasta que

tengamos un total de 8. Esto nos lleva al neón. A medida que añadimos estos electrones estamos también aumentando Z , por lo cual la distribución electrónica total se comprime más y más cerca del núcleo y la energía de los estados $2p$ disminuye. Cuando llegamos al neón la energía de ionización ha aumentado a 21.6 voltios. El neón no suelta fácilmente un electrón. Tampoco hay más lugares de energía baja que se puedan ocupar, por lo que no tratará de agarrar un electrón adicional. El neón es inerte químicamente. El flúor, por otra parte, *sí* tiene un lugar vacío donde un electrón puede caer en un estado de energía baja y, por lo tanto, es muy activo en reacciones químicas.

Na hasta Ar

En el sodio el electrón undécimo debe empezar una nueva capa —yendo a un estado $3s$ —. El nivel de energía de este estado es mucho mayor; la energía de ionización decrece, y el sodio es químicamente activo. Desde el sodio hasta el argón, los estados s y p con $n = 3$ están ocupados exactamente en la misma secuencia que desde el litio al neón. Las configuraciones angulares de los electrones en las capas externas no completas tienen la misma secuencia y el aumento de las energías de ionización es bastante similar. Pueden con ello ver por qué las propiedades químicas se repiten con el aumento del número atómico. El magnesio actúa químicamente en forma muy parecida al berilio, el silicio al carbono y el cloro al flúor. El argón es inerte como el neón.

Pueden haber notado que hay una leve particularidad en la secuencia de las energías de ionización entre el litio y el neón, y una semejante entre el sodio y el argón. El último electrón está ligado al átomo de oxígeno algo menos de lo que podíamos esperar. Y al azufre le ocurre algo parecido. ¿A qué se debe eso? Podemos entenderlo si colocamos solamente un poquito de los efectos de interacción entre electrones individuales. Consideren lo que sucede cuando ponemos el primer electrón $2p$ en el átomo de boro. Tiene seis posibilidades —tres estados posibles p , cada uno con dos espines. Imaginen que el electrón va al estado $m = 0$ con espín hacia arriba, que también hemos llamado estado “ z ” porque abraza al eje z . Pues bien, ¿qué sucederá en el carbono? Ahora hay dos electrones $2p$. Si uno de ellos va al estado “ z ”, ¿adónde irá el segundo? Tendrá una energía menor si se aleja del primer electrón, lo que puede hacer yendo, digamos, al estado “ x ” de la capa $2p$ (recuerden que este estado es justamente una combinación lineal de los estados $m = +1$ y $m = -1$). Luego cuando pasamos al nitrógeno, los tres electrones $2p$ tendrán la energía más baja de repulsión mutua si van a cada una de las configuraciones “ x ”, “ y ” y “ z ”. Para el oxígeno, sin embargo, no hay más esperanza. El cuarto electrón debe ir a uno de los estados completos —con espín opuesto—. Es fuertemente repelido por el electrón que ya se encuentra en ese estado, por lo que su energía no será tan baja como sin la repulsión y será más fácil de arrancar. Esto explica la brecha en la secuencia de energías de ligadura que aparece entre el nitrógeno y el oxígeno y entre el fósforo y el silicio.

K hasta Zn

Después del argón, a primera vista pensarían que los nuevos electrones empezarían a llenar los estados $3d$. Pero no es así. Como describimos anteriormente —e ilustramos en la figura 19-8— los estados de momentum angular más alto son empujados hacia energías más altas. Para cuando llegamos a los estados $3d$ han sido empujados a una energía un poco por encima de la energía del estado $4s$. Por eso el último electrón del potasio va al estado $4s$. Después que esta capa está completa (con dos electrones) en el calcio, comienzan a llenarse los estados $3d$ para el escandio, el titanio y el vanadio.

Las energías de los estados $3d$ y $4s$ están tan juntas que efectos pequeños pueden inclinar el equilibrio a un lado o a otro. Al colocar cuatro electrones en los estados $3d$, su repulsión aumenta la energía de los estados $4s$ en forma tal que su energía está un poco por encima de la energía $3d$, por lo que se traspasa un electrón. Para el cromo no obtenemos una combinación 4, 2, como esperaríamos,

sino una combinación 5, 1. El nuevo electrón añadido para obtener el manganeso completa de nuevo la capa 4s y los estados de la capa 3d se ocupan entonces uno a uno hasta que llegamos al cobre.

Como la capa más externa del manganeso, el hierro, el cobalto y el níquel tiene las mismas configuraciones, todos ellos tienden a tener propiedades químicas similares (este efecto es mucho más pronunciado en los elementos de las tierras raras que tienen la misma capa externa, pero una capa interna que se llena progresivamente y que influye mucho menos en sus propiedades químicas).

En el cobre se roba un electrón de la capa 4s, completándose así la capa 3d. La energía de la combinación 10, 1 es, sin embargo, tan próxima a la configuración 9, 2 para el cobre que sólo la presencia de un átomo cercano puede hacer perder el equilibrio. Por esta razón, los dos últimos electrones del cobre son casi equivalentes y el cobre puede tener una valencia de 1 o 2 (algunas veces actúa como si sus electrones estuviesen en la combinación 9, 2). Cosas análogas suceden en otros lugares y explican el hecho de que otros metales, como por ejemplo, el hierro, se combinen químicamente con 1 o con 2 valencias. En el zinc, las capas 3d y 4s se completan de una vez por todas.

Ga hasta Kr

Desde el galio hasta el criptón la secuencia sigue normalmente de nuevo, completando la capa 4p. Las capas externas, las energías, y las propiedades químicas repiten el modelo desde el boro hasta el neón y desde el aluminio hasta el argón.

El criptón, así como el argón y el neón, se conocen como gases nobles. Los tres son “inertes” químicamente. Esto sólo significa que, teniendo capas completas de energía relativamente baja, hay pocas situaciones que tengan una ventaja energética para unirse con otros elementos en una combinación simple. Tener una capa completa no es bastante. El berilio y el magnesio tienen capas *s* completas, pero la energía de estas capas es demasiado alta para que los haga estables. Del mismo modo, se hubiera esperado otro elemento “noble” en el níquel, si la energía de la capa 3d hubiese sido más baja (o la 4s más alta). Por otra parte, el criptón no es completamente inerte; formará con cloro un compuesto débilmente ligado.

Como nuestro ejemplo ha apuntado la mayoría de las características principales de la tabla periódica, detenemos nuestro examen en el elemento 36 —¡aún hay setenta y pico más!

Queríamos agregar otro punto: no solamente podemos entender las valencias hasta cierto modo, sino que también podemos decir algo acerca de las propiedades direccionales de los enlaces químicos. Tomen un átomo como el oxígeno que tiene cuatro electrones 2p. Los tres primeros van a los estados “*x*”, “*y*” y “*z*” y el cuarto doblará uno de estos estados, dejando dos —digamos “*x*” y “*y*”— vacantes. Consideren, entonces, lo que sucede en H₂O. Cada uno de los dos hidrógenos está deseando compartir un electrón con el oxígeno, ayudándole a completar una capa. Estos electrones tenderán a ir hacia los huecos “*x*” y “*y*”. Por lo tanto, la molécula de agua debería tener los dos átomos de hidrógeno formando un ángulo recto respecto al centro del oxígeno. El ángulo es realmente 105°. Hasta podemos entender por qué el ángulo es mayor de 90°. Al compartir sus electrones, los hidrógenos acaban con una carga positiva neta. La repulsión eléctrica “fuerza” las funciones de onda y empuja el ángulo a 105°. Lo mismo ocurre en H₂S. Pero como el azufre es un átomo más grande, los dos átomos de hidrógeno están más separados, hay menos repulsión, y el ángulo es empujado solamente a alrededor de 93°. El selenio es aún más grande, por lo que en H₂Se el ángulo es casi de 90°.

Podemos utilizar los mismos razonamientos para entender la geometría del amoníaco, H₃N. El nitrógeno tiene sitio para tres electrones 2p más, para los tipos de estado “*x*”, “*y*” y “*z*”. Los tres hidrógenos se deberían unir uno a otro en ángulo recto. Los ángulos resultan un poquito mayores de 90° —nuevamente debido a la repulsión eléctrica—, pero al menos vemos por qué la molécula de H₃N no es plana. Los ángulos en el fosfeno, H₃P, son próximos a 90° y en H₃As

aún más próximos. Supusimos que NH_3 no era plana cuando la describimos como un sistema de dos estados. Y el no ser plana es lo que hace posible el máser de amoníaco. Ahora vemos que también de esa forma se pueden entender con nuestra mecánica cuántica.

La ecuación de Schrödinger ha sido uno de los grandes triunfos de la física. Al darnos la clave de la maquinaria interna de la estructura atómica nos ha dado una explicación de los espectros atómicos, de la química y de la naturaleza de la materia.

Operadores

20-1 Operaciones y operadores

Todo lo que hemos hecho hasta ahora en la mecánica cuántica se puede hacer con álgebra ordinaria, aunque de vez en cuando les mostramos algunas maneras especiales de escribir cantidades y ecuaciones cuánticas. Ahora nos gustaría hablar un poco más sobre algunas maneras matemáticas interesantes y útiles de describir cosas cuánticas. Hay varias maneras de enfocar la mecánica cuántica y la mayoría de los libros usa un enfoque diferente del que hemos adoptado. Cuando sigan estudiando y lean otros libros puede ser que no vean inmediatamente la conexión entre lo que encontrarán en ellos y lo que hemos estado haciendo. Aunque podríamos obtener algunos resultados útiles, el fin principal de este capítulo es darles a conocer algunas maneras nuevas de escribir la misma física. Conociéndolas, estarán en condiciones de comprender mejor lo que otros dicen. Cuando se comenzó a desarrollar la mecánica clásica, los físicos siempre escribían las ecuaciones en términos de las componentes x , y y z . Luego vino alguien y señaló que se podía simplificar mucho la escritura inventando la notación vectorial. Es verdad que cuando llega el momento de calcular algo, a menudo hay que volver a convertir los vectores en sus componentes. Pero generalmente es mucho más fácil ver qué está pasando cuando trabajan con vectores y además es más fácil hacer muchos de los cálculos. En la mecánica cuántica pudimos escribir muchas cosas de un modo más simple usando el concepto de “vector de estado”. Por supuesto que el vector de estado $|\psi\rangle$ no tiene nada que ver con vectores geométricos en tres dimensiones; éste es un símbolo abstracto que *representa un estado físico*, identificado con el “rótulo” o “nombre” ψ . El concepto es útil porque empleando estos símbolos es posible escribir las leyes de la mecánica cuántica como ecuaciones algebraicas. Por ejemplo, nuestra ley fundamental de que se puede construir cualquier estado como combinación lineal de estados base se escribe en la forma

$$|\psi\rangle = \sum_i C_i |i\rangle, \quad (20.1)$$

donde las C_i son un conjunto de números ordinarios (complejos) —las amplitudes $C_i = \langle i|\psi\rangle$ — mientras que $|1\rangle, |2\rangle, |3\rangle$, etc., representan los estados base de alguna *representación* o base.

Si toman un estado físico y le hacen algo —tal como rotarlo o esperar durante un tiempo Δt — obtienen un estado diferente. Decimos “realizar una operación sobre un estado produce un nuevo estado”. Podemos expresar la misma idea con una ecuación:

$$|\phi\rangle = \hat{A}|\psi\rangle. \quad (20.2)$$

Una operación sobre un estado produce otro estado. El *operador* $\hat{A}$ representa una operación determinada. Cuando se realiza esta operación sobre cualquier estado, por ejemplo $|\psi\rangle$, éste produce algún otro estado $|\phi\rangle$.

¿Qué significa la ecuación (20.2)? La *definimos* de esta manera. Si multiplican la ecuación por $\langle i|$ y desarrollan $|\psi\rangle$ conforme a la ecuación (20.1), obtienen

$$\langle i|\phi\rangle = \sum_j \langle i|\hat{A}|j\rangle \langle j|\psi\rangle. \quad (20.3)$$

(Los estados $|j\rangle$ son del mismo conjunto que $|i\rangle$.) Ésta es simplemente una ecuación algebraica. Los números $\langle i|\phi\rangle$ dan la cantidad de cada estado base que encontrarán en $|\phi\rangle$, y están dados como superposición lineal de las amplitudes

20-1 Operaciones y operadores

20-2 Valores medios de la energía

20-3 Valor medio de la energía de un átomo

20-4 El operador posición

20-5 El operador momentum

20-6 Momentum angular

20-7 Variación temporal de valores medios

$\langle j | \psi \rangle$ de que encuentren $|\psi\rangle$ en cada estado base. Los números $\langle i | \hat{A} | j \rangle$ son simplemente los coeficientes que dicen cuánto $\langle j | \psi \rangle$ va a cada sumando. El operador $\hat{A}$ está descrito numéricamente por el conjunto de números, o “matriz”,

$$A_{ij} \equiv \langle i | \hat{A} | j \rangle. \quad (20.4)$$

Así pues, la ecuación (20.2) es una manera elegante de escribir (20.3). En realidad es un poco más que eso; interviene algo más. En la ecuación (20.2) no hacemos ninguna referencia a un conjunto de estados base. La ecuación (20.3) es una imagen de la ecuación (20.2) en términos de cierto conjunto de estados base. Pero como saben, pueden utilizar el conjunto que deseen. Y esta idea está implícita en la ecuación (20.2). La manera operacional de escribir les evita hacer una elección determinada. Por supuesto, cuando quieran hacer una aplicación específica tendrán que elegir *algún* conjunto. Cuando hagan su elección, usen la ecuación (20.3). Así, pues, la ecuación *operacional* (20.2) es una manera más abstracta de escribir la ecuación *algebraica* (20.3). Es análogo a la diferencia entre escribir

$$\mathbf{c} = \mathbf{a} \times \mathbf{b}$$

en vez de

$$c_x = a_y b_z - a_z b_y,$$

$$c_y = a_z b_x - a_x b_z,$$

$$c_z = a_x b_y - a_y b_x.$$

La primera es mucho más práctica. Sin embargo, cuando quieran *resultados*, tendrán que dar a la postre las componentes respecto a un conjunto de ejes. Análogamente, si quieren poder decir qué entienden realmente por $\hat{A}$, tendrán que estar listos para dar la matriz A_{ij} en función de *algún* conjunto de estados base. Mientras tengan en mente algún conjunto $|i\rangle$, la ecuación (20.2) significa exactamente lo mismo que la ecuación (20.3) (deben recordar además que una vez que conocen una matriz para un conjunto particular de estados base, siempre pueden calcular la correspondiente matriz en cualquier otra base. Pueden transformar la matriz de una “representación” a otra).

La ecuación operacional (20.2) también permite una nueva manera de pensar. Si imaginamos algún operador $\hat{A}$, podemos usarlo con cualquier estado $|\psi\rangle$ para crear un nuevo estado $\hat{A}|\psi\rangle$. A veces, el “estado” que obtenemos de esta manera puede ser muy raro —puede que no represente ninguna situación *física* de las que probablemente encontremos en la naturaleza (por ejemplo, podemos obtener un estado que no está normalizado para representar un electrón)—. En otras palabras, a veces podemos obtener “estados” que son matemáticamente artificiales. Esos “estados” artificiales pueden ser útiles de todas maneras, quizá como paso intermedio en algún cálculo.

Ya les hemos mostrado muchos ejemplos de operadores cuánticos. Hemos tenido el operador de rotación $\hat{R}_y(\theta)$ que toma un estado $|\psi\rangle$ y produce otro, que es el viejo visto en un sistema de coordenadas rotado. Hemos tenido el operador de paridad (o de inversión) $\hat{P}$, que hace un nuevo estado invirtiendo todas las coordenadas. Hemos tenido los operadores $\hat{\sigma}_x$, $\hat{\sigma}_y$ y $\hat{\sigma}_z$ para partículas de espín un medio.

En el capítulo 17 se definió el operador $\hat{J}_z$ en función del operador de rotación para un ángulo pequeño ϵ :

$$\hat{R}_z(\epsilon) = 1 + \frac{i}{\hbar} \epsilon \hat{J}_z. \quad (20.5)$$

Por supuesto que esto significa simplemente

$$\hat{R}_z(\epsilon) |\psi\rangle = |\psi\rangle + \frac{i}{\hbar} \epsilon \hat{J}_z |\psi\rangle. \quad (20.6)$$

En este ejemplo, $\hat{J}_z |\psi\rangle$ es $\hbar/i\epsilon$ por el estado que obtienen si rotan $|\psi\rangle$ en un ángulo pequeño ϵ y luego restan el estado original. Representa un “estado” que es la *diferencia* de dos estados.

Un ejemplo más. Tuvimos un operador $\hat{p}_x$ —llamado operador momentum (componente x) definido en una ecuación como la (20.6)—. Si $\hat{D}_x(L)$ es el operador que desplaza un estado según x la distancia L , $\hat{p}_x$ está definido por

$$\hat{D}_x(\delta) = 1 + \frac{i}{\hbar} \delta \hat{p}_x, \quad (20.7)$$

donde δ es un desplazamiento pequeño. Desplazando el estado $|\psi\rangle$ según x una distancia pequeña δ se obtiene un nuevo estado $|\psi'\rangle$. Estamos diciendo que este nuevo estado es el viejo más un pequeño pedazo nuevo:

$$\frac{i}{\hbar} \delta \hat{p}_x |\psi\rangle.$$

Los operadores de que estamos hablando actúan sobre un vector de estado tal como $|\psi\rangle$, el cual es una descripción abstracta de una situación física. Son completamente diferentes de los operadores *algebraicos* que actúan sobre funciones matemáticas. Por ejemplo, d/dx es un “operador” que actúa sobre $f(x)$ transformándola en una nueva función $f'(x) = df/dx$. Otro ejemplo es el operador algebraico ∇^2 . Pueden ver por qué se usa la misma palabra en ambos casos, pero deben recordar que las dos clases de operadores son diferentes. Un operador cuántico $\hat{A}$ no actúa sobre una función algebraica, sino sobre un vector de estado tal como $|\psi\rangle$. En la mecánica cuántica se usan ambas clases de operadores y a menudo en clases similares de ecuaciones, como verán dentro de poco. Al aprender el tema por primera vez siempre es bueno tener presente la distinción. Más adelante, cuando estén más familiarizados con el tema, encontrarán que no tiene tanta importancia mantener una distinción nítida entre las dos clases de operadores. Por cierto, encontrarán que la mayoría de los libros usa generalmente la misma notación para ambas!

Continuemos ahora y examinemos algunas cosas útiles que se pueden hacer con operadores. Pero antes hagamos un comentario especial. Supongan que tenemos un operador, $\hat{A}$ cuya matriz en cierta base es $A_{ij} \equiv \langle i | \hat{A} | j \rangle$. La amplitud de que el estado $\hat{A}|\psi\rangle$ también esté en algún otro estado $|\phi\rangle$ es $\langle \phi | \hat{A} | \psi \rangle$. ¿Tiene algún significado la compleja conjugada de esta amplitud? Ustedes deben ser capaces de demostrar que

$$\langle \phi | \hat{A} | \psi \rangle^* = \langle \psi | \hat{A}^\dagger | \phi \rangle, \quad (20.8)$$

donde $\hat{A}^\dagger$ (lean “A cruz”) es un operador cuyos elementos de matriz son

$$A_{ij}^\dagger = (A_{ji})^*. \quad (20.9)$$

Para obtener el elemento i, j de $\hat{A}^\dagger$ van al elemento i, j de $\hat{A}$ (se invierten los índices) y toman su complejo conjugado. La amplitud de que el estado $\hat{A}^\dagger|\phi\rangle$ esté en $|\psi\rangle$ es el complejo conjugado de la amplitud de que $\hat{A}|\psi\rangle$ esté en $|\phi\rangle$. El operador $\hat{A}^\dagger$ se llama “adjunto hermitiano” de $\hat{A}$. Muchos operadores importantes de la mecánica cuántica tienen la propiedad especial de que cuando toman el adjunto hermitiano obtienen otra vez el mismo operador. Un operador $\hat{B}$ como éste satisface entonces

$$\hat{B}^\dagger = \hat{B},$$

y se llama operador “autoadjunto” o “hermítico”.

20-2 Valores medios de la energía

Hasta ahora les hemos recordado principalmente cosas que ya saben. Nos gustaría estudiar un punto nuevo. ¿Cómo hallarían el *valor medio* de la energía de un sistema —un átomo, por ejemplo—? Si un átomo está en un estado determinado de energía definida y miden la energía, encontrarán cierta energía E . Si repiten la medida sobre cada uno de los átomos de una serie, elegidos de modo que estén en el mismo estado, todas las medidas darán E y el “valor medio” de sus medidas será E , por supuesto.

Ahora bien, ¿qué ocurre si hacen la medida sobre un estado $|\psi\rangle$ que *no* es un estado estacionario? Como el sistema no tiene una energía definida, una medida dará una energía, la misma medida sobre otro átomo en el mismo estado dará una energía diferente y así sucesivamente. ¿Qué obtienen para el valor medio de toda una serie de medidas de energía?

Podemos contestar la pregunta proyectando el estado $|\psi\rangle$ sobre un conjunto de estados de energía definida. Para recordarles que éste es un conjunto base especial, llamaremos $|\eta_i\rangle$ a los estados. Cada estado $|\eta_i\rangle$ tiene una energía definida E_i . En esta representación,

$$|\psi\rangle = \sum_i C_i |\eta_i\rangle. \quad (20.10)$$

Al hacer una medida de energía y obtener un número E_i , han encontrado que el sistema estaba en el estado η_i . Pero en cada medida pueden obtener un número diferente. A veces obtienen E_1 , a veces E_2 , a veces E_3 y así sucesivamente. La *probabilidad* de que observen la energía E_1 es justamente la probabilidad de encontrar el sistema en el estado $|\eta_1\rangle$, que naturalmente no es sino el cuadrado del módulo de la amplitud $C_1 = \langle\eta_1|\psi\rangle$. La probabilidad de encontrar cada energía posible E_i es

$$P_i = |C_i|^2. \quad (20.11)$$

¿Cómo están relacionadas estas probabilidades con el valor medio de toda una sucesión de medidas de energía? Imaginen que obtenemos una serie de medidas como ésta: $E_1, E_7, E_{11}, E_9, E_1, E_{10}, E_7, E_2, E_3, E_9, E_6, E_4$, etc. Digamos que continuamos hasta mil medidas. Cuando terminamos, sumamos todas las energías y dividimos entre mil. Esto es lo que entendemos por valor medio. Hay también un atajo para sumar todos los números. Pueden contar cuántas veces obtienen E_1 , digamos que N_1 , y luego cuentan el número de veces que obtienen E_2 , llámenlo N_2 y así sucesivamente. La suma de todas las energías es simplemente

$$N_1 E_1 + N_2 E_2 + N_3 E_3 + \cdots = \sum_i N_i E_i.$$

El valor medio de la energía es esta suma dividida entre el número total de medidas, que es sencillamente la suma de todos los N_i y que podemos llamar N ;

$$E_{av} = \frac{\sum_i N_i E_i}{N}. \quad (20.12)$$

Ya estamos llegando. Lo que *entendemos* por probabilidad de que ocurra algo es justamente el número de veces que esperamos que ocurra dividido entre el número total de intentos. El cociente N_i/N debe ser —para N grande— muy cercano a P_i , probabilidad de encontrar el estado $|\eta_i\rangle$, aunque no será exactamente P_i debido a las fluctuaciones estadísticas. Designemos con $\langle E \rangle_{av}$ el valor medio predicho (o “esperado”) de la energía; podemos decir entonces que

$$\langle E \rangle_{av} = \sum_i P_i E_i. \quad (20.13)$$

El mismo razonamiento sirve para cualquier medida. El valor medio de una cantidad medida A debe ser igual a

$$\langle A \rangle_{av} = \sum_i P_i A_i,$$

donde los A_i son los diversos valores posibles de la cantidad observada y P_i es la probabilidad de obtener ese valor.

Volvamos a nuestro estado cuántico $|\psi\rangle$. El valor medio de su energía es

$$\langle E \rangle_{av} = \sum_i |C_i|^2 E_i = \sum_i C_i^* C_i E_i. \quad (20.14)$$

¡Y miren ahora esta trampa! Primero escribimos la suma en la forma

$$\sum_i \langle \psi | \eta_i \rangle E_i \langle \eta_i | \psi \rangle. \quad (20.15)$$

Luego tratamos el $\langle \psi |$ de la izquierda como un “factor” común. Podemos sacar este factor de la suma y escribirla en la forma

$$\langle \psi | \left\{ \sum_i | \eta_i \rangle E_i \langle \eta_i | \psi \rangle \right\}.$$

Esta expresión es de la forma

$$\langle \psi | \phi \rangle,$$

donde $|\phi\rangle$ es un estado “prefabricado” que se define por

$$|\phi\rangle = \sum_i | \eta_i \rangle E_i \langle \eta_i | \psi \rangle. \quad (20.16)$$

En otras palabras, es el estado que obtienen si toman una cantidad $E_i \langle \eta_i | \psi \rangle$ de cada estado base $|\eta_i\rangle$.

Recuerden ahora lo que entendemos por los estados $|\eta_i\rangle$. Se supone que son los estados estacionarios, con lo cual queremos decir que para cada uno,

$$\hat{H} | \eta_i \rangle = E_i | \eta_i \rangle.$$

Como E_i es simplemente un número, el segundo miembro es lo mismo que $|\eta_i\rangle E_i$ y la suma en la ecuación (20.16) es lo mismo que

$$\sum_i \hat{H} | \eta_i \rangle \langle \eta_i | \psi \rangle.$$

Ahora i aparece en la famosa combinación que da la unidad, por lo que

$$\sum_i \hat{H} | \eta_i \rangle \langle \eta_i | \psi \rangle = \hat{H} \sum_i | \eta_i \rangle \langle \eta_i | \psi \rangle = \hat{H} | \psi \rangle.$$

¡Magia! La ecuación (20.16) es lo mismo que

$$|\phi\rangle = \hat{H} | \psi \rangle. \quad (20.17)$$

El valor medio de la energía en el estado $|\psi\rangle$ se puede escribir muy elegantemente en la forma

$$\langle E \rangle_{\text{av}} = \langle \psi | \hat{H} | \psi \rangle. \quad (20.18)$$

Para obtener el valor medio de la energía se operan con $\hat{H}$ sobre $|\psi\rangle$ y luego multiplican por $\langle \psi |$. Es un resultado simple.

Nuestra nueva fórmula para el valor medio de la energía no es únicamente elegante. También es útil, porque ahora no necesitamos decir nada sobre cualquier conjunto particular de estados base. Ni siquiera tenemos que conocer todos los niveles de energía posibles. Cuando vayamos a calcular necesitaremos describir nuestro estado en términos de *algún* conjunto de estados base, pero si conocemos la matriz hamiltoniana H_{ij} para *ese* conjunto podemos obtener el valor medio de la energía. La ecuación (20.18) dice que para *cualquier* conjunto de estados base $|i\rangle$, se puede calcular el valor medio de la energía con

$$\langle E \rangle_{\text{av}} = \sum_{ij} \langle \psi | i \rangle \langle i | \hat{H} | j \rangle \langle j | \psi \rangle, \quad (20.19)$$

donde las amplitudes $\langle i | \hat{H} | j \rangle$ son simplemente los elementos de la matriz H_{ij} .

Verifiquemos este resultado para el caso especial en que los estados $|i\rangle$ sean los estados de energía definida. Así, $\hat{H} | j \rangle = E_j | j \rangle$, por lo que $\langle i | \hat{H} | j \rangle = E_j \delta_{ij}$ y

$$\langle E \rangle_{\text{av}} = \sum_{ij} \langle \psi | i \rangle E_j \delta_{ij} \langle j | \psi \rangle = \sum_i E_i \langle \psi | i \rangle \langle i | \psi \rangle,$$

que es lo correcto.

Incidentalmente, se puede extender la ecuación (20.19) a otras medidas físicas que puedan expresarse con un operador. Por ejemplo, $\hat{L}_z$ es el operador de la componente z del momentum angular $\mathbf{L}$. El valor medio de la componente z en el estado $|\psi\rangle$ es

$$\langle L_z \rangle_{\text{av}} = \langle \psi | \hat{L}_z | \psi \rangle.$$

Una manera de demostrarlo es considerar alguna situación en la que la energía sea proporcional al momentum angular. Entonces el razonamiento se hace igual que antes.

En resumen, si un observable físico A está relacionado con un operador cuántico apropiado $\hat{A}$, el valor medio de A en el estado $|\psi\rangle$ está dado por

$$\langle A \rangle_{\text{av}} = \langle \psi | \hat{A} | \psi \rangle. \quad (20.20)$$

Con esto queremos decir

$$\langle A \rangle_{\text{av}} = \langle \psi | \phi \rangle, \quad (20.21)$$

con

$$|\phi\rangle = \hat{A} |\psi\rangle. \quad (20.22)$$

20-3 Valor medio de la energía de un átomo

Supongan que queremos el valor medio de la energía de un átomo en un estado descrito por una función de onda $\psi(\mathbf{r})$; ¿cómo hallarlo? Consideremos primero una situación unidimensional con un estado $|\psi\rangle$ definido por la amplitud $\langle x | \psi \rangle = \psi(x)$. Estamos buscando el caso especial de la ecuación (20.19) aplicada a la representación de coordenadas. Siguiendo nuestro procedimiento habitual, reemplazamos los estados $|i\rangle$ y $|j\rangle$ por $|x\rangle$ y $|x'\rangle$ y cambiamos por integrales. Obtenemos

$$\langle E \rangle_{\text{av}} = \int \int \langle \psi | x \rangle \langle x | \hat{H} | x' \rangle \langle x' | \psi \rangle dx dx'. \quad (20.23)$$

Si lo deseamos, se puede escribir esta integral de la manera siguiente:

$$\int \langle \psi | x \rangle \langle x | \phi \rangle dx, \quad (20.24)$$

con

$$\langle x | \phi \rangle = \int \langle x | \hat{H} | x' \rangle \langle x' | \psi \rangle dx'. \quad (20.25)$$

La integral sobre x' en (20.25) es la misma que teníamos en el capítulo 16 —véanse las ecuaciones (16.50) y (16.52)— y es igual a

$$-\frac{\hbar^2}{2m} \frac{d^2}{dx^2} \psi(x) + V(x)\psi(x).$$

Por lo tanto, podemos escribir

$$\langle x | \phi \rangle = \left\{ -\frac{\hbar^2}{2m} \frac{d^2}{dx^2} + V(x) \right\} \psi(x). \quad (20.26)$$

Recuerden que $\langle \psi | x \rangle = \langle x | \psi \rangle^* = \psi^*(x)$; usando esta igualdad, se puede escribir el valor medio (20.23) de la energía en la forma

$$\langle E \rangle_{\text{av}} = \int \psi^*(x) \left\{ -\frac{\hbar^2}{2m} \frac{d^2}{dx^2} + V(x) \right\} \psi(x) dx. \quad (20.27)$$

Dada una función de onda $\psi(x)$, pueden obtener el valor medio de la energía haciendo esta integral. Pueden empezar a ver cómo vamos y volvemos entre los conceptos de vector de estado y de función de onda.

En la ecuación (20.27) la cantidad entre llaves es un operador *algebraico*.† Lo escribiremos $\hat{\mathcal{H}}$,

$$\hat{\mathcal{H}} = -\frac{\hbar^2}{2m} \frac{d^2}{dx^2} + V(x).$$

† El “operador” $V(x)$ significa “multiplicar por $V(x)$ ”.

En esta notación la ecuación (20.23) es

$$\langle E \rangle_{\text{av}} = \int \psi^*(x) \hat{\mathcal{H}} \psi(x) dx. \quad (20.28)$$

El operador algebraico $\hat{\mathcal{H}}$ definido aquí, no es, naturalmente, idéntico al operador cuántico $\hat{H}$. El nuevo operador actúa sobre una función de la posición $\psi(x) = \langle x | \psi \rangle$ para dar una nueva función de x , $\phi(x) = \langle x | \phi \rangle$; mientras que $\hat{H}$ opera sobre un vector de estado $|\psi\rangle$ para dar otro vector de estado $|\phi\rangle$, sin que esto implique la representación de coordenadas o cualquier otra representación determinada. Ni aun en la representación de coordenadas $\hat{\mathcal{H}}$ es estrictamente igual a $\hat{H}$. Si eligiéramos trabajar en la representación de coordenadas, interpretaríamos $\hat{H}$ en términos de una matriz $\langle x | \hat{H} | x' \rangle$ que de alguna manera depende de los dos “índices” x y x' ; es decir, esperamos —conforme a la ecuación (20.25)— que $\langle x | \phi \rangle$ esté relacionada con todas las amplitudes $\langle x | \psi \rangle$ por medio de una integración. Por otra parte, encontramos que $\hat{\mathcal{H}}$ es un operador diferencial. Ya hemos obtenido en la sección 16-5 la conexión entre $\langle x | \hat{H} | x' \rangle$ y el operador algebraico $\hat{\mathcal{H}}$.

Debemos hacer una limitación a nuestros resultados. Hemos estado suponiendo que la amplitud $\psi(x) = \langle x | \psi \rangle$ está normalizada. Con esto queremos decir que se ha elegido la escala de modo que

$$\int |\psi(x)|^2 dx = 1;$$

así, la probabilidad de encontrar el electrón en *cualquier parte* es uno. Si escogieran trabajar con una $\psi(x)$ sin normalizar tendrían que escribir

$$\langle E \rangle_{\text{av}} = \frac{\int \psi^*(x) \hat{\mathcal{H}} \psi(x) dx}{\int \psi^*(x) \psi(x) dx}. \quad (20.29)$$

Es lo mismo.

Observen la similitud de forma entre la ecuación (20.28) y la ecuación (20.18). Estas dos maneras de escribir el mismo resultado aparecen a menudo cuando trabajan en la representación x . Pueden pasar de la primera forma a la segunda con cualquier $\hat{A}$ que sea un operador *local*, o sea, un operador que en la integral

$$\int \langle x | \hat{A} | x' \rangle \langle x' | \psi \rangle dx'$$

se puede escribir como $\hat{A}\psi(x)$, donde $\hat{A}$ es un operador algebraico diferencial. Sin embargo, hay operadores para los que esto no vale. En ese caso tienen que trabajar con las ecuaciones básicas (20.21) y (20.22).

Pueden extender fácilmente el desarrollo a tres dimensiones. El resultado es†

$$\langle E \rangle_{\text{av}} = \int \psi^*(\mathbf{r}) \hat{\mathcal{H}} \psi(\mathbf{r}) d\text{Vol}, \quad (20.30)$$

con

$$\hat{\mathcal{H}} = -\frac{\hbar^2}{2m} \nabla^2 + V(\mathbf{r}), \quad (20.31)$$

y entendiendo implícitamente que

$$\int |\psi|^2 d\text{Vol} = 1. \quad (20.32)$$

Se pueden extender las mismas ecuaciones a sistemas de varios electrones, de un modo bastante evidente, pero no nos molestaremos en escribir los resultados.

Con la ecuación (20.30) podemos calcular la energía media de un estado atómico aun sin conocer sus niveles de energía. Todo lo que necesitamos es la función de onda. Es una ley importante. Les hablaremos de una aplicación

† Escribimos $d\text{Vol}$ para el elemento de volumen. Por supuesto, es simplemente $dx dy dz$, y la integral es entre $-\infty$ y $+\infty$ para las tres coordenadas.

interesante. Supongan que quieren saber cuál es la energía del estado fundamental de un sistema —del átomo de helio, por ejemplo—, pero es muy difícil resolver la ecuación de Schrödinger para la función de onda porque hay demasiadas variables. Supongan, sin embargo, que imaginan la función de onda —elijan la función que les guste— y calculen el valor medio de la energía. Es decir, usan la ecuación (20.29) —generalizada a tres dimensiones— para hallar cuál sería el valor medio de la energía si el átomo estuviera verdaderamente en el estado descrito por esa función de onda. Con certeza, esta energía será mayor que la energía del estado fundamental que es la energía más baja posible que puede tener el átomo.† Tomen ahora otra función y calculen el valor medio de la energía. Si es menor que el de la primera elección, se están acercando a la verdadera energía del estado fundamental. Si siguen probando todo tipo de estados artificiales, podrán obtener energías cada vez más bajas que se acercan más y más a la energía del estado fundamental. Si son astutos, probarán funciones que tengan algunos parámetros ajustables. Al calcular la energía quedará expresada en función de estos parámetros. Al variar los parámetros para dar la energía más baja posible, están probando a la vez toda una clase de funciones. En un momento dado encontrarán que es más y más difícil obtener energías más bajas y comenzarán a convencerse que están bastante cerca de la energía más baja posible. Precisamente, se ha resuelto el átomo de helio de esta manera —no resolviendo una ecuación diferencial, sino fabricando una función especial con un montón de parámetros ajustables que finalmente se eligen para dar el valor más bajo posible del valor medio de la energía—.

20-4 El operador posición

¿Cuál es el valor medio de la posición de un electrón en un átomo? En un estado cualquiera $|\psi\rangle$, ¿cuál es el valor medio de la coordenada x ? Trabajaremos en una dimensión y dejaremos que ustedes extiendan los conceptos a tres dimensiones o a sistemas con más de una partícula. Tenemos un estado descrito por $\psi(x)$ y nos mantenemos midiendo x una y otra vez. ¿Cuál es el promedio? Es

$$\int xP(x) dx,$$

donde $P(x)$ es la probabilidad de encontrar el electrón en un pequeño elemento dx alrededor de x . Supongan que la densidad de probabilidad $P(x)$ varía con x como muestra la figura 20-1. Es más probable encontrar el electrón cerca del pico de la curva. El valor medio de x también está en algún lugar cerca del pico. De hecho, es justamente el centro de gravedad del área bajo la curva.

Hemos visto anteriormente que $P(x)$ es simplemente $|\psi(x)|^2 = \psi^*(x)\psi(x)$, por lo que podemos escribir el valor medio de x en la forma

$$\langle x \rangle_{av} = \int \psi^*(x)x\psi(x) dx. \quad (20.33)$$

Nuestra ecuación para $\langle x \rangle_{av}$ tiene la misma forma que la ecuación (20.28). En el valor medio de la energía, el operador energía $\hat{\mathcal{H}}$ aparece entre las dos ψ ; en el valor medio de la posición está simplemente x (si lo desean pueden considerar que x es el operador algebraico “multiplicar por x ”). Podemos llevar el paralelismo aún más allá, expresando la posición media en una forma que corresponda a la ecuación (20.18). Supongan que escribimos

$$\langle x \rangle_{av} = \langle \psi | \alpha \rangle \quad (20.34)$$

con

$$| \alpha \rangle = \hat{x} | \psi \rangle, \quad (20.35)$$

† También pueden verlo de esta otra manera. Cualquier función (es decir, estado) que elijan se puede escribir como combinación lineal de los estados base que son estados de energía definida. Como en esta combinación hay estados mayores de energía mezclados con el estado más bajo de energía, la energía media será mayor que la energía del estado fundamental.

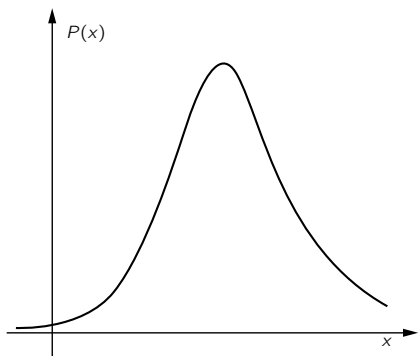


Figura 20-1. Una curva de densidad de probabilidad que representa una partícula localizada.

y luego vemos si podemos encontrar el operador $\hat{x}$ que genera el estado $|\alpha\rangle$, lo cual hará que la ecuación (20.34) concuerde con la ecuación (20.33). Es decir, tenemos que hallar un $|\alpha\rangle$ tal que

$$\langle\psi|\alpha\rangle = \langle x\rangle_{\text{av}} = \int \langle\psi|x\rangle x \langle x|\psi\rangle dx. \quad (20.36)$$

Primero desarrollemos $\langle\psi|\alpha\rangle$ en la representación x . Es

$$\langle\psi|\alpha\rangle = \int \langle\psi|x\rangle \langle x|\alpha\rangle dx. \quad (20.37)$$

Comparen ahora las integrales que aparecen en las dos últimas ecuaciones. Ven que en la representación x

$$\langle x|\alpha\rangle = x \langle x|\psi\rangle. \quad (20.38)$$

Operar sobre $|\psi\rangle$ con $\hat{x}$ para obtener $|\alpha\rangle$ es equivalente a multiplicar $\psi(x) = \langle x|\psi\rangle$ por x para obtener $\alpha(x) = \langle x|\alpha\rangle$. Tenemos una definición de $\hat{x}$ en la representación de coordenadas.[†]

[No nos hemos molestado en tratar de obtener la representación x de la matriz del operador $\hat{x}$. Si son ambiciosos pueden tratar de demostrar que

$$\langle x|\hat{x}|x'\rangle = x \delta(x - x'). \quad (20.39)$$

Luego pueden obtener el siguiente resultado divertido:

$$\hat{x}|x\rangle = x|x\rangle. \quad (20.40)$$

El operador $\hat{x}$ tiene la propiedad interesante de que cuando actúa sobre los estados base $|x\rangle$ es equivalente a multiplicar por x .]

¿Quieren conocer el valor medio de x^2 ? Es

$$\langle x^2\rangle_{\text{av}} = \int \psi^*(x) x^2 \psi(x) dx. \quad (20.41)$$

Y si lo prefieren, pueden escribir

$$\langle x^2\rangle_{\text{av}} = \langle\psi|\alpha'\rangle$$

con

$$|\alpha'\rangle = \hat{x}^2|\psi\rangle. \quad (20.42)$$

Por $\hat{x}^2$ entendemos $\hat{x}\hat{x}$ —se usa los dos operadores uno después de otro—. Con la segunda forma pueden calcular $\langle x^2\rangle_{\text{av}}$ empleando cualquier representación (cualquier base) que deseen. Si quieren el valor medio de x^n o de cualquier polinomio en x , pueden ver cómo obtenerlo.

20-5 El operador momentum

Ahora nos gustaría calcular el *momentum* medio de un electrón —de nuevo nos limitaremos a una dimensión—. Sea $P(p)dp$ la probabilidad de que una medida dé un momentum entre p y $p + dp$. Entonces

$$\langle p\rangle_{\text{av}} = \int p P(p) dp. \quad (20.43)$$

Ahora bien, sea $\langle p|\psi\rangle$ la amplitud de que el estado $|\psi\rangle$ esté en un estado $|p\rangle$ de un momentum definido. Es la misma amplitud que llamamos $\langle \text{mom } p|\psi\rangle$ en

[†] La ecuación (20.38) *no* significa que $|\alpha\rangle = x|\psi\rangle$. No pueden “sacar factor común” $\langle x|$, porque el multiplicador x delante de $\langle x|\psi\rangle$ es un número diferente para cada estado $\langle x|$. Es el valor de la coordenada del electrón en el estado $|x\rangle$. Véase la ecuación (20.40).

la sección 16-3 y es una función de p así como $\langle x | \psi \rangle$ es una función de x . Allí escogimos la normalización de la amplitud de modo que

$$P(p) = \frac{1}{2\pi\hbar} |\langle p | \psi \rangle|^2. \quad (20.44)$$

Tenemos entonces,

$$\langle p \rangle_{\text{av}} = \int \langle \psi | p \rangle p \langle p | \psi \rangle \frac{dp}{2\pi\hbar}. \quad (20.45)$$

La forma es completamente similar a la que teníamos para $\langle x \rangle_{\text{av}}$.

Si quieren, podemos hacer el mismo juego que hicimos con $\langle x \rangle_{\text{av}}$. En primer lugar, podemos escribir la integral anterior en la forma

$$\int \langle \psi | p \rangle \langle p | \beta \rangle \frac{dp}{2\pi\hbar}. \quad (20.46)$$

Ahora deben reconocer en esta ecuación la forma desarrollada de la amplitud $\langle \psi | \beta \rangle$ —desarrollada en los estados base de momentum definido—. Según la ecuación (20.45) el estado $|\beta\rangle$ está definido *en la representación del momentum* por

$$\langle p | \beta \rangle = p \langle p | \psi \rangle \quad (20.47)$$

Es decir, ahora podemos escribir

$$\langle p \rangle_{\text{av}} = \langle \psi | \beta \rangle \quad (20.48)$$

con

$$|\beta\rangle = \hat{p} |\psi\rangle, \quad (20.49)$$

donde el operador $\hat{p}$ está definido en la representación p mediante la ecuación (20.47).

[También aquí, si lo desean, pueden demostrar que la forma matricial de $\hat{p}$ es

$$\langle p | \hat{p} | p' \rangle = p \delta(p - p'), \quad (20.50)$$

y que

$$\hat{p} | p \rangle = p | p \rangle. \quad (20.51)$$

Es como en el caso de x .]

Ahora surge una pregunta interesante. Podemos escribir $\langle p \rangle_{\text{av}}$ tal como en las ecuaciones (20.45) y (20.48), y conocemos el significado del operador $\hat{p}$ en la *representación de momentum*. ¿Pero cómo interpretaríamos $\hat{p}$ en la representación de *coordenadas*? Eso es lo que necesitaremos saber si tenemos una función de onda $\psi(x)$ y queremos calcular el valor medio de su momentum. Aclaremos lo que queremos decir. Si comenzamos diciendo que $\langle p \rangle_{\text{av}}$ está dado por la ecuación (20.48), podemos desarrollar esa ecuación en la representación p y obtener la ecuación (20.46) de vuelta. Si se nos da la descripción del estado en la representación p —o sea, la amplitud $\langle p | \psi \rangle$, que es una función algebraica del momentum p — podemos obtener $\langle p | \beta \rangle$ de la ecuación (20.47) y calcular la integral. Y ahora surge la pregunta: ¿qué hacemos si nos dan una descripción del estado en la representación x , o sea, la función de onda $\psi(x) = \langle x | \psi \rangle$?

Empecemos desarrollando la ecuación (20.48) en la representación x . Es

$$\langle p \rangle_{\text{av}} = \int \langle \psi | x \rangle \langle x | \beta \rangle dx. \quad (20.52)$$

No obstante, necesitamos saber cómo es el estado $|\beta\rangle$ en la representación x . Si lo podemos encontrar, podremos formar la integral. Así, pues, nuestro problema es hallar la función $\beta(x) = \langle x | \beta \rangle$.

Podemos hallarla de la siguiente manera. En la sección 16-3 vimos cómo estaba relacionada $\langle p | \beta \rangle$ con $\langle x | \beta \rangle$. Según la ecuación (16.24),

$$\langle p | \beta \rangle = \int e^{-ipx/\hbar} \langle x | \beta \rangle dx. \quad (20.53)$$

Si conocemos $\langle p | \beta \rangle$ podemos despejar $\langle x | \beta \rangle$ de esta ecuación. Lo que queremos es, naturalmente, expresar el resultado de alguna manera en términos de $\psi(x) = \langle x | \psi \rangle$, que suponemos se conoce. Supongan que partimos de la ecuación (20.47) y usamos la ecuación (16.24) de nuevo para escribir

$$\langle p | \beta \rangle = p \langle p | \psi \rangle = p \int e^{-ipx/\hbar} \psi(x) dx. \quad (20.54)$$

Como la integral es sobre x , podemos poner p dentro de la integral y escribir

$$\langle p | \beta \rangle = \int e^{-ipx/\hbar} p \psi(x) dx. \quad (20.55)$$

Compárenla con (20.53). Dirían que $\langle x | \beta \rangle$ es igual a $p\psi(x)$. ¡No y no! La función de onda $\langle x | \beta \rangle = \beta(x)$ sólo puede depender de x y no de p . Ése es todo el problema.

Sin embargo, un tipo ingenioso descubrió que la integral de (20.55) se podía integrar por partes. La derivada de $e^{-ipx/\hbar}$ respecto a x es $(-i/\hbar)pe^{-ipx/\hbar}$, por lo que la integral de (20.55) es equivalente a

$$-\frac{\hbar}{i} \int \frac{d}{dx} (e^{-ipx/\hbar}) \psi(x) dx.$$

Integrando por partes se obtiene

$$-\frac{\hbar}{i} [e^{-ipx/\hbar} \psi(x)]_{-\infty}^{+\infty} + \frac{\hbar}{i} \int e^{-ipx/\hbar} \frac{d\psi}{dx} dx.$$

Mientras nos limitemos a estados ligados, de modo que $\psi(x)$ tienda a cero para $x = \pm\infty$, el corchete es cero y tenemos

$$\langle p | \beta \rangle = \frac{\hbar}{i} \int e^{-ipx/\hbar} \frac{d\psi}{dx} dx. \quad (20.56)$$

Ahora comparen este resultado con la ecuación (20.53). Ven que

$$\langle x | \beta \rangle = \frac{\hbar}{i} \frac{d}{dx} \psi(x). \quad (20.57)$$

Tenemos la parte que nos faltaba para completar la ecuación (20.52). La respuesta es

$$\langle p \rangle_{\text{av}} = \int \psi^*(x) \frac{\hbar}{i} \frac{d}{dx} \psi(x) dx. \quad (20.58)$$

Hemos encontrado cuál es el aspecto de la ecuación (20.48) en la representación de coordenadas.

Ahora podrían darse cuenta del modelo tan interesante que se desarrolla. Cuando preguntamos por el valor medio de la energía en el estado $|\psi\rangle$ dijimos que era

$$\langle E \rangle_{\text{av}} = \langle \psi | \phi \rangle, \quad \text{con} \quad |\phi\rangle = \hat{H} |\psi\rangle.$$

En el mundo de las coordenadas, lo mismo se escribe en la forma

$$\langle E \rangle_{\text{av}} = \int \psi^*(x) \phi(x) dx \quad \text{con} \quad \phi(x) = \hat{\mathcal{H}} \psi(x).$$

Aquí $\hat{\mathcal{H}}$ es un operador *algebraico* que actúa sobre una función de x . Cuando preguntamos por el valor medio de x encontramos que también se podía escribir

$$\langle x \rangle_{\text{av}} = \langle \psi | \alpha \rangle, \quad \text{con} \quad |\alpha\rangle = \hat{x} |\psi\rangle.$$

En el mundo de las coordenadas las ecuaciones correspondientes son

$$\langle x \rangle_{\text{av}} = \int \psi^*(x) \alpha(x) dx, \quad \text{con} \quad \alpha(x) = x\psi(x).$$

Cuando nos preguntamos por el valor medio de p , escribimos

$$\langle p \rangle_{\text{av}} = \langle \psi | \hat{p} | \psi \rangle, \quad \text{con} \quad | \beta \rangle = \hat{p} | \psi \rangle.$$

En el mundo de las coordenadas las ecuaciones equivalentes eran

$$\langle p \rangle_{\text{av}} = \int \psi^*(x) \beta(x) dx, \quad \text{con} \quad \beta(x) = \frac{\hbar}{i} \frac{d}{dx} \psi(x).$$

En cada uno de los tres ejemplos empezamos con el estado $|\psi\rangle$ y producimos otro estado (hipotético) mediante un operador *cuántico*. En la representación de coordenadas generamos la función de onda correspondiente operando sobre la función de onda $\psi(x)$ con un operador *algebraico*. Hay las siguientes correspondencias unívocas (para problemas unidimensionales):

$$\begin{aligned} \hat{H} &\rightarrow \hat{\mathcal{H}} = -\frac{\hbar^2}{2m} \frac{d^2}{dx^2} + V(x), \\ \hat{x} &\rightarrow x, \\ \hat{p}_x &\rightarrow \hat{\mathcal{P}}_x = \frac{\hbar}{i} \frac{\partial}{\partial x}. \end{aligned} \tag{20.59}$$

En esta lista hemos introducido el símbolo $\hat{\mathcal{P}}_x$ para el operador algebraico $(\hbar/i)\partial/\partial x$:

$$\hat{\mathcal{P}}_x = \frac{\hbar}{i} \frac{\partial}{\partial x}, \tag{20.60}$$

y hemos introducido el subíndice x en $\hat{\mathcal{P}}$ para recordarles que hemos estado trabajando únicamente con la componente x del momentum.

Fácilmente pueden extender los resultados a tres dimensiones. Para las otras componentes del momentum,

$$\begin{aligned} \hat{p}_y &\rightarrow \hat{\mathcal{P}}_y = \frac{\hbar}{i} \frac{\partial}{\partial y}, \\ \hat{p}_z &\rightarrow \hat{\mathcal{P}}_z = \frac{\hbar}{i} \frac{\partial}{\partial z}. \end{aligned}$$

Si quieren, hasta pueden imaginar un operador del momentum *vectorial* y escribir

$$\hat{\mathbf{p}} \rightarrow \hat{\mathcal{P}} = \frac{\hbar}{i} \left(\mathbf{e}_x \frac{\partial}{\partial x} + \mathbf{e}_y \frac{\partial}{\partial y} + \mathbf{e}_z \frac{\partial}{\partial z} \right),$$

Tabla 20-1

<i>Cantidad física</i>	<i>Operador</i>	<i>Formulario de coordenadas</i>
Energía	$\hat{H}$	$\hat{\mathcal{H}} = -\frac{\hbar^2}{2m} \nabla^2 + V(\mathbf{r})$
Posición	$\hat{x}$	x
	$\hat{y}$	y
	$\hat{z}$	z
Momentum	$\hat{p}_x$	$\hat{\mathcal{P}}_x = \frac{\hbar}{i} \frac{\partial}{\partial x}$
	$\hat{p}_y$	$\hat{\mathcal{P}}_y = \frac{\hbar}{i} \frac{\partial}{\partial y}$
	$\hat{p}_z$	$\hat{\mathcal{P}}_z = \frac{\hbar}{i} \frac{\partial}{\partial z}$

donde $\mathbf{e}_x$, $\mathbf{e}_y$ y $\mathbf{e}_z$ son vectores en las tres direcciones. Es mucho más elegante escribir

$$\hat{\mathbf{p}} \rightarrow \hat{\mathbf{P}} = \frac{\hbar}{i} \nabla. \quad (20.61)$$

Nuestro resultado general es que por lo menos para algunos operadores cuánticos existen los correspondientes operadores algebraicos en la representación de coordenadas. En la tabla 20-1 resumimos los resultados obtenidos hasta ahora —extendidos a tres dimensiones—. Para cada operador tenemos las dos formas equivalentes:†

$$|\phi\rangle = \hat{A}|\psi\rangle \quad (20.62)$$

o

$$\phi(\mathbf{r}) = \hat{A}\psi(\mathbf{r}). \quad (20.63)$$

Unos pocos ejemplos ilustrarán la utilidad de estas ideas. El primero consiste en señalar la relación entre $\hat{\mathcal{P}}$ y $\hat{\mathcal{H}}$. Si usamos $\hat{\mathcal{P}}_x$ dos veces obtenemos

$$\hat{\mathcal{P}}_x \hat{\mathcal{P}}_x = -\hbar^2 \frac{\partial^2}{\partial x^2}.$$

Esto significa que podemos escribir la igualdad

$$\hat{\mathcal{H}} = \frac{1}{2m} \{ \hat{\mathcal{P}}_x \hat{\mathcal{P}}_x + \hat{\mathcal{P}}_y \hat{\mathcal{P}}_y + \hat{\mathcal{P}}_z \hat{\mathcal{P}}_z \} + V(\mathbf{r}).$$

O sea, en notación vectorial,

$$\hat{\mathcal{H}} = \frac{1}{2m} \hat{\mathbf{P}} \cdot \hat{\mathbf{P}} + V(\mathbf{r}). \quad (20.64)$$

(En un operador algebraico, cualquier término sin el símbolo (^) de operador significa simplemente una multiplicación.) Esta ecuación es muy interesante porque es fácil de recordar si no han olvidado su física clásica. Cualquiera sabe que la energía es (en forma no relativista) la energía cinética $p^2/2m$ más la energía potencial, y $\hat{\mathcal{H}}$ es el operador de la energía total.

Este resultado ha impresionado tanto a los físicos que éstos han tratado de enseñar a los estudiantes toda la física clásica antes que la mecánica cuántica (¡nosotros pensamos en forma diferente!). Pero a menudo estos paralelismos llevan a confusión. Sólo para mencionar una cosa: cuando tienen operadores, el *orden* de los diversos factores es importante; pero esto no vale para los factores de una ecuación clásica.

En el capítulo 17 definimos un operador $\hat{p}_x$ en términos del operador desplazamiento $\hat{D}_x$, mediante [véase la ecuación (17.27)]

$$|\psi'\rangle = \hat{D}_x(\delta)|\psi\rangle = \left(1 + \frac{i}{\hbar} \hat{p}_x \delta\right)|\psi\rangle. \quad (20.65)$$

donde δ es un desplazamiento *pequeño*. Deberíamos mostrarles que esto es equivalente a nuestra nueva definición. Según lo que acabamos de calcular, esta ecuación debe significar lo mismo que

$$\psi'(x) = \psi(x) + \frac{\partial \psi}{\partial x} \delta.$$

Pero el segundo miembro es simplemente el desarrollo de Taylor de $\psi(x + \delta)$ que es por cierto lo que obtienen si desplazan el estado hacia la izquierda en la cantidad δ (o corren las coordenadas hacia la derecha en la misma cantidad). ¡Nuestras dos definiciones de $\hat{p}$ concuerdan!

Usemos esto para demostrar algo más. Supongan que tenemos un enjambre de partículas que designamos 1, 2, 3, ..., formando algún sistema complicado (para

† En muchos libros se usa el mismo símbolo para $\hat{A}$ y $\hat{\mathcal{A}}$, porque ambos representan lo mismo en física y porque no es conveniente tener que escribir diferentes tipos de letra. De ordinario pueden deducir cuál es cuál por el contexto.

simplificar, nos limitaremos a una dimensión). La función de onda que describe el estado es una función de todas las coordenadas $x_1, x_2, x_3, \dots$. Podemos escribirla $\psi(x_1, x_2, x_3, \dots)$. Ahora desplacen el sistema (hacia la izquierda) la cantidad δ . La nueva función de onda

$$\psi'(x_1, x_2, x_3, \dots) = \psi(x_1 + \delta, x_2 + \delta, x_3 + \delta, \dots)$$

se puede escribir en la forma

$$\psi'(x_1, x_2, x_3, \dots) = \psi(x_1, x_2, x_3, \dots) + \left\{ \delta \frac{\partial \psi}{\partial x_1} + \delta \frac{\partial \psi}{\partial x_2} + \delta \frac{\partial \psi}{\partial x_3} + \dots \right\}. \quad (20.66)$$

De acuerdo con la ecuación (20.65) el operador momentum del estado $|\psi\rangle$ (llamémoslo momentum *total*) es igual a

$$\hat{P}_{\text{total}} = \frac{\hbar}{i} \left\{ \frac{\partial}{\partial x_1} + \frac{\partial}{\partial x_2} + \frac{\partial}{\partial x_3} + \dots \right\}.$$

Pero esto es justamente lo mismo que

$$\hat{P}_{\text{total}} = \hat{P}_{x_1} + \hat{P}_{x_2} + \hat{P}_{x_3} + \dots. \quad (20.67)$$

Los operadores momentum obedecen la regla de que el momentum total es la suma de los momenta de todas las partes. Todo ensambla muy bien y hay gran concordancia en mucho de lo que hemos estado diciendo.

20-6 Momentum angular

Para divertirnos, examinemos otra operación —la operación momentum angular orbital—. En el capítulo 17 definimos un operador $\hat{J}_z$ en función de $\hat{R}_z(\phi)$, el operador de una rotación en un ángulo ϕ alrededor del eje z . Consideraremos aquí un sistema descrito simplemente por una sola función de onda $\psi(\mathbf{r})$, que es función de las coordenadas únicamente, y que no toma en cuenta que el electrón puede tener su espín hacia arriba o hacia abajo. Es decir, por el momento queremos dejar de lado el momentum angular *intrínseco* y pensar únicamente en la parte *orbital*. Para que la distinción sea clara, llamaremos $\hat{L}_z$ al operador orbital y lo definiremos en función del operador de rotación en un ángulo infinitesimal ϵ mediante

$$\hat{R}_z(\epsilon) |\psi\rangle = \left(1 + \frac{i}{\hbar} \epsilon \hat{L}_z \right) |\psi\rangle.$$

(Recuerden que esta definición sólo sirve para un estado $|\psi\rangle$ que no tiene variables de espín, sino que sólo depende de las coordenadas $\mathbf{r} = x, y, z$.) Si miramos $|\psi\rangle$ en un nuevo sistema de coordenadas rotado en un ángulo pequeño ϵ del eje z , vemos el nuevo estado

$$|\psi'\rangle = \hat{R}_z(\epsilon) |\psi\rangle.$$

Si elegimos la representación de coordenadas para representar el estado $|\psi\rangle$ —es decir, con su función de onda $\psi(\mathbf{r})$ — sería de esperar que pudiéramos escribir

$$\psi'(\mathbf{r}) = \left(1 + \frac{i}{\hbar} \epsilon \hat{L}_z \right) \psi(\mathbf{r}). \quad (20.68)$$

¿Qué es $\hat{L}_z$? Bueno, un punto P de coordenadas x, y en el *nuevo* sistema (en realidad x', y' , pero no escribiremos las primas) estaba antes en $x - \epsilon y$ y $y + \epsilon x$, como pueden ver en la figura 20-2. Como la amplitud de que el electrón esté en P no cambia al rotar las coordenadas, podemos escribir

$$\psi'(x, y, z) = \psi(x - \epsilon y, y + \epsilon x, z) = \psi(x, y, z) - \epsilon y \frac{\partial \psi}{\partial x} + \epsilon x \frac{\partial \psi}{\partial y}$$

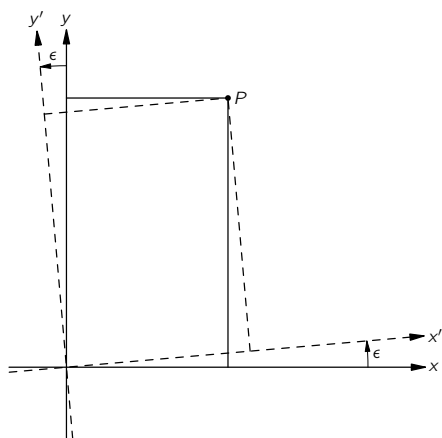


Figura 20-2. Rotación de los ejes alrededor del eje z por un ángulo pequeño ϵ .

(recordando que ϵ es un ángulo pequeño). Esto significa que

$$\hat{\mathcal{L}}_z = \frac{\hbar}{i} \left(x \frac{\partial}{\partial y} - y \frac{\partial}{\partial x} \right). \quad (20.69)$$

Ésta es nuestra respuesta. Pero observen: es equivalente a

$$\hat{\mathcal{L}}_z = x\hat{\mathcal{P}}_y - y\hat{\mathcal{P}}_x. \quad (20.70)$$

Volviendo a nuestros operadores cuánticos, podemos escribir

$$\hat{L}_z = x\hat{p}_y - y\hat{p}_x. \quad (20.71)$$

Esta fórmula es fácil de recordar porque se parece a la fórmula familiar de la mecánica clásica; es la componente z de

$$\mathbf{L} = \mathbf{r} \times \mathbf{p}. \quad (20.72)$$

Uno de los aspectos divertidos de este negocio de operadores es que muchas ecuaciones clásicas se pueden trasladar a una forma cuántica. ¿Cuáles no? Es mejor que haya algunas que no se pueda porque, si se pudiera con todas, no habría nada diferente en la mecánica cuántica. No habría ningún aspecto físico nuevo. Aquí hay una ecuación que es diferente. En la física clásica

$$xp_x - p_x x = 0.$$

¿Cómo es en la mecánica cuántica?

$$\hat{x}\hat{p}_x - \hat{p}_x\hat{x} = ?$$

Hallémoslo en la representación x . Para saber lo que estamos haciendo introduciremos una función de onda $\psi(x)$. Tenemos

$$x\hat{\mathcal{P}}_x\psi(x) - \hat{\mathcal{P}}_xx\psi(x),$$

o sea

$$x \frac{\hbar}{i} \frac{\partial}{\partial x} \psi(x) - \frac{\hbar}{i} \frac{\partial}{\partial x} x\psi(x).$$

Recuerden ahora que las derivadas operan sobre todo lo que está a la derecha. Obtenemos

$$x \frac{\hbar}{i} \frac{\partial \psi}{\partial x} - \frac{\hbar}{i} \psi(x) - \frac{\hbar}{i} x \frac{\partial \psi}{\partial x} = -\frac{\hbar}{i} \psi(x). \quad (20.73)$$

La respuesta *no* es cero. Toda la operación es simplemente equivalente a multiplicar por $-\hbar/i$:

$$\hat{x}\hat{p}_x - \hat{p}_x\hat{x} = -\frac{\hbar}{i}. \quad (20.74)$$

Si la constante de Planck fuera cero, los resultados clásicos y los cuánticos serían iguales ¡y no habría mecánica cuántica que aprender!

Incidentalmente, si cuando se toma dos operadores $\hat{A}$ y $\hat{B}$ cualesquiera así:

$$\hat{A}\hat{B} - \hat{B}\hat{A},$$

no da cero, decimos que “los operadores no conmutan”. Y una ecuación tal como la (20.74) se llama “regla de conmutación”. Pueden ver que la regla de conmutación para p_x y y es

$$\hat{p}_x\hat{y} - \hat{y}\hat{p}_x = 0.$$

Hay otra regla de conmutación muy importante que se refiere a los momenta angulares. Es

$$\hat{L}_x\hat{L}_y - \hat{L}_y\hat{L}_x = i\hbar\hat{L}_z. \quad (20.75)$$

Demostrándola por sí mismos podrán adquirir práctica con los operadores $\hat{x}$ y $\hat{p}$.

Es interesante observar que también en la física clásica puede haber operadores que no conmutan. Ya hemos visto esto al hablar de rotaciones en el espacio. Si rotan algo, tal como un libro, 90° alrededor de x y luego 90° alrededor de y , obtienen algo diferente de rotar primero 90° alrededor de y y luego 90° alrededor de x . De hecho, es precisamente esta propiedad la que da lugar a la ecuación (20.75).

20-7 Variación temporal de valores medios

Ahora queremos mostrarles algo distinto. ¿Cómo varían los valores medios en el tiempo? Supongan por el momento que tenemos un operador $\hat{A}$ que no contiene el tiempo de modo evidente. Entendemos un operador tal como $\hat{x}$ o $\hat{p}$. [Excluimos cosas tales como el operador de un potencial externo que está variando en el tiempo, tal como $V(x, t)$.] Supongan ahora que calculamos $\langle A \rangle_{\text{av}}$ en un estado $|\psi\rangle$, lo cual es

$$\langle A \rangle_{\text{av}} = \langle \psi | \hat{A} | \psi \rangle. \quad (20.76)$$

¿Cómo dependerá $\langle A \rangle_{\text{av}}$ del tiempo? ¿Por qué debe depender? Una razón podría ser que el operador mismo dependiera explícitamente del tiempo —por ejemplo, que tuviera que ver con un potencial $V(x, t)$ que varía en el tiempo—. Pero aunque el operador no dependa de t , digamos, por ejemplo, el operador $\hat{A} = \hat{x}$, el valor medio correspondiente puede depender del tiempo. Es seguro que la posición media de una partícula podría estarse moviendo. ¿Cómo se origina ese movimiento en la ecuación (20.76) si $\hat{A}$ no tiene dependencia temporal? Bueno, el estado $|\psi\rangle$ podría estar cambiando con el tiempo. Para los estados no estacionarios, a menudo hemos mostrado una dependencia temporal explícita escribiendo el estado en la forma $|\psi(t)\rangle$. Queremos demostrar que la derivada de $\langle A \rangle_{\text{av}}$ respecto al tiempo está dada por un nuevo operador que llamaremos $\hat{\dot{A}}$. Recuerden que $\hat{A}$ es un operador, por lo cual poner un punto encima de la A no significa aquí tomar la derivada respecto al tiempo, sino que es solamente una manera de escribir un *nuevo* operador $\hat{\dot{A}}$ que está definido por

$$\frac{d}{dt} \langle A \rangle_{\text{av}} = \langle \psi | \hat{\dot{A}} | \psi \rangle. \quad (20.77)$$

Nuestro problema es encontrar el operador $\hat{\dot{A}}$.

Primero: sabemos que la derivada de un estado respecto al tiempo está dada por el hamiltoniano. Específicamente

$$i\hbar \frac{d}{dt} |\psi(t)\rangle = \hat{H} |\psi(t)\rangle. \quad (20.78)$$

Esto es sólo una manera abstracta de escribir nuestra definición original del hamiltoniano:

$$i\hbar \frac{dC_i}{dt} = \sum_j H_{ij} C_j. \quad (20.79)$$

Si tomamos el complejo conjugado de la ecuación (20.78), es equivalente a

$$-i\hbar \frac{d}{dt} \langle \psi(t) | = \langle \psi(t) | \hat{H}. \quad (20.80)$$

Segundo: observen lo que ocurre si derivamos la ecuación (20.76) respecto al tiempo. Como cada ψ depende de t , tenemos

$$\frac{d}{dt} \langle A \rangle_{\text{av}} = \left(\frac{d}{dt} \langle \psi | \right) \hat{A} | \psi \rangle + \langle \psi | \hat{A} \left(\frac{d}{dt} | \psi \rangle \right). \quad (20.81)$$

Finalmente: usando las ecuaciones (20.78) y (20.80) para remplazar las derivadas, obtenemos

$$\frac{d}{dt} \langle A \rangle_{\text{av}} = \frac{i}{\hbar} \{ \langle \psi | \hat{H} \hat{A} | \psi \rangle - \langle \psi | \hat{A} \hat{H} | \psi \rangle \}.$$

Esta ecuación es lo mismo que

$$\frac{d}{dt} \langle A \rangle_{\text{av}} = \frac{i}{\hbar} \langle \psi | \hat{H} \hat{A} - \hat{A} \hat{H} | \psi \rangle.$$

Comparando esta ecuación con (20.77), ven que

$$\hat{\dot{A}} = \frac{i}{\hbar} (\hat{H} \hat{A} - \hat{A} \hat{H}). \quad (20.82)$$

Éste es nuestro enunciado interesante; es válido para cualquier operador $\hat{A}$.

Incidentalmente, si el operador $\hat{A}$ dependiera *él mismo* del tiempo, habríamos tenido

$$\hat{A} = \frac{i}{\hbar} (\hat{H}\hat{A} - \hat{A}\hat{H}) + \frac{\partial \hat{A}}{\partial t}. \quad (20.83)$$

Probemos la ecuación (20.82) en cierto ejemplo para ver si realmente tiene sentido. Por ejemplo, ¿qué operador corresponde a $\hat{x}$? Decimos que debe ser

$$\hat{x} = \frac{i}{\hbar} (\hat{H}\hat{x} - \hat{x}\hat{H}). \quad (20.84)$$

¿Qué es esto? Una manera de hallarlo es trabajar en la representación de coordenadas usando el operador algebraico $\mathcal{H}$. En esta representación el conmutador es

$$\hat{\mathcal{H}}x - x\hat{\mathcal{H}} = \left\{ -\frac{\hbar^2}{2m} \frac{d^2}{dx^2} + V(x) \right\} x - x \left\{ -\frac{\hbar^2}{2m} \frac{d^2}{dx^2} + V(x) \right\}.$$

Si operan con éste sobre cualquier función de onda $\psi(x)$ y donde puedan calcular todas las derivadas, llegan después de un poco de trabajo a

$$-\frac{\hbar^2}{m} \frac{d\psi}{dx}.$$

Pero esto es precisamente lo mismo que

$$-i \frac{\hbar}{m} \hat{p}_x \psi,$$

por lo que encontramos

$$\hat{H}\hat{x} - \hat{x}\hat{H} = -i \frac{\hbar}{m} \hat{p}_x \quad (20.85)$$

o sea que

$$\hat{x} = \frac{\hat{p}_x}{m}. \quad (20.86)$$

Un lindo resultado. Significa que si el valor medio de x está variando en el tiempo, la velocidad del centro de gravedad es igual al valor medio del momentum dividido entre m . Exactamente como en la mecánica clásica.

Otro ejemplo. ¿Cuál es la derivada respecto al tiempo, del valor medio del momentum de un estado? El mismo juego. Su operador es

$$\hat{p} = \frac{i}{\hbar} (\hat{H}\hat{p} - \hat{p}\hat{H}). \quad (20.87)$$

Nuevamente pueden calcularlo en la representación x . Recuerden que $\hat{p}$ se convierte en d/dx , lo cual significa que estarán derivando la energía potencial V (en $\mathcal{H}$) —pero sólo en el segundo término—. Resulta que es el único que no se anula y encuentran

$$\hat{\mathcal{H}}\hat{p} - \hat{p}\hat{\mathcal{H}} = i\hbar \frac{dV}{dx}$$

o sea

$$\hat{p} = -\frac{dV}{dx}. \quad (20.88)$$

Nuevamente el resultado clásico. El segundo miembro es la fuerza y así ¡hemos deducido la ley de Newton! Pero recuerden: estas leyes son para los *operadores* que dan cantidades *medias*. No describen lo que pasa en detalle dentro de un átomo.

La mecánica cuántica tiene la diferencia esencial de que $\hat{p}\hat{x}$ no es igual a $\hat{x}\hat{p}$. Difieren un poquitito —en el número pequeño $i\hbar$ —. Pero todas las complicaciones maravillosas de la interferencia, las ondas y todo ello resultan del hecho insignificante de que $\hat{x}\hat{p} - \hat{p}\hat{x}$ no es del todo cero.

La historia de esta idea también es interesante. En un periodo de unos pocos meses en 1926, Heisenberg y Schrödinger encontraron independientemente leyes

correctas para describir la mecánica del átomo. Schrödinger inventó su función de onda $\psi(x)$ y encontró su ecuación. Heisenberg, por otra parte, encontró que se podía describir la naturaleza con ecuaciones clásicas, exceptuando que $xp - px$ debería ser igual a $i\hbar$, y pudo hacer que esto se cumpliera definiéndolos en términos de tipos especiales de matrices. En nuestro lenguaje, estaba usando la representación de energía con sus matrices. El álgebra de matrices de Heisenberg tanto como la ecuación diferencial de Schrödinger explicaron el átomo de hidrógeno. Pocos meses más tarde Schrödinger pudo demostrar que las dos teorías eran equivalentes —como hemos visto aquí—. Pero las dos formas matemáticas diferentes de la mecánica cuántica fueron descubiertas independientemente.

La ecuación de Schrödinger en el contexto clásico: un seminario sobre superconductividad

21-1 La ecuación de Schrödinger en un campo magnético

Damos esta lección solamente como distracción. Me gustaría hacerlo en un estilo un poco diferente —para ver cómo resulta—. No es una parte del curso —en el sentido de que no suponemos que sea un esfuerzo de última hora para enseñarles algo nuevo—. Pensamos, más bien, en un seminario o en un informe de investigación para una audiencia más avanzada, para gente ya cultivada en la mecánica cuántica. La diferencia principal entre un seminario y una lección regular es que el director de un seminario no realiza todos los pasos o resuelve toda el álgebra. Dice: “Si hacen esto y esto, resulta esto”, en lugar de mostrarles todos los detalles. Así, en toda esta lección describiré las ideas a lo largo de todo el proceso, pero sólo les daré los *resultados* de los cálculos. Deben darse cuenta de que no suponemos que van a entenderlo todo inmediatamente, pero crean (más o menos) que encontrarían ese resultado si efectuasen todos los pasos.

Aparte de eso, éste es un tema del que *quiero* hablar. Es reciente y moderno y sería una charla perfectamente legítima para darla en un seminario de investigación. El tema es: la ecuación de Schrödinger en un marco clásico —el caso de la superconductividad—.

Comúnmente, la función de onda que aparece en la ecuación de Schrödinger se aplica solamente a una o a dos partículas. Y la función de onda misma no es algo que tenga un significado clásico —al contrario del campo eléctrico, o del potencial vectorial, o de cosas por el estilo—. La función de onda para una sola partícula *es* un “campo” —en el sentido de que es una función de la posición— pero generalmente no tiene un significado clásico. No obstante, hay algunos casos en los cuales una función de onda cuántica *sí* tiene significado clásico, y son los que me gustaría tomar. El comportamiento cuántico peculiar de la materia a escala microscópica por lo general no se hace sentir a escala macroscópica, excepto en la forma normal en que produce las leyes de Newton —las leyes de la llamada mecánica clásica—. Pero hay un cierto número de situaciones en las cuales las peculiaridades de la mecánica cuántica pueden aparecer de una manera especial a escala macroscópica.

A bajas temperaturas, donde la energía de un sistema ha disminuido muchísimo, interviene solamente un pequeñísimo número de estados cercanos al estado fundamental, en vez de un gran número de ellos. En estas circunstancias, el carácter cuántico del estado fundamental puede aparecer a escala macroscópica. El propósito de esta lección es mostrar una conexión entre la mecánica cuántica y los efectos macroscópicos —no es la discusión común del modo en que la mecánica cuántica reproduce en promedio la mecánica newtoniana, sino una situación especial en la cual la mecánica cuántica producirá sus propios efectos característicos a una escala grande o “macroscópica”—.

Comenzaré recordándoles algunas de las propiedades de la ecuación de Schrödinger.† Quiero describir el comportamiento de una partícula en un campo magnético usando la ecuación de Schrödinger, ya que los fenómenos de superconductividad están interrelacionados con campos magnéticos. Un campo magnético externo se describe mediante un potencial vectorial y el problema es: ¿cuáles

21-1 La ecuación de Schrödinger en un campo magnético

21-2 Ecuación de continuidad para probabilidades

21-3 Dos clases de momentum

21-4 El significado de la función de onda

21-5 Superconductividad

21-6 El efecto Meissner

21-7 Cuantificación del flujo

21-8 Dinámica de la superconductividad

21-9 Juntura de Josephson

† Realmente no se las estoy recordando, ya que no les he mostrado antes algunas de estas ecuaciones; pero recuerden el espíritu de este seminario.

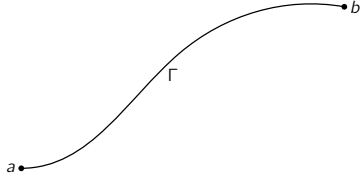


Figura 21-1. La amplitud de ir de a a b a lo largo de la trayectoria Γ es proporcional a $\exp[(iq/\hbar) \int_a^b \mathbf{A} \cdot d\mathbf{s}]$.

son las leyes de la mecánica cuántica en un potencial vectorial? El principio que describe el comportamiento de la mecánica cuántica en un potencial vectorial es muy simple. La amplitud de que una partícula vaya de un punto a otro a través de un cierto camino cuando existe un campo, es igual a la amplitud de que siga el mismo camino en ausencia del campo, multiplicada por la exponencial de la integral de línea del potencial vectorial, por la carga eléctrica dividida entre la constante de Planck¹ (véase la figura 21-1):

$$\langle b | a \rangle_{\text{en } \mathbf{A}} = \langle b | a \rangle_{A=0} \cdot \exp \left[\frac{iq}{\hbar} \int_a^b \mathbf{A} \cdot d\mathbf{s} \right]. \quad (21.1)$$

Éste es un enunciado básico de la mecánica cuántica.

Ahora bien, sin el potencial vectorial la ecuación de Schrödinger de una partícula cargada (no relativista, sin espín) es

$$-\frac{\hbar}{i} \frac{\partial \psi}{\partial t} = \hat{\mathcal{H}} \psi = \frac{1}{2m} \left(\frac{\hbar}{i} \nabla \right) \cdot \left(\frac{\hbar}{i} \nabla \right) \psi + q\phi\psi, \quad (21.2)$$

donde ϕ es el potencial eléctrico y, por lo tanto, $q\phi$ es la energía potencial.[†] La ecuación (21.1) es equivalente al enunciado de que en un campo magnético los gradientes que aparecen en el hamiltoniano se remplazan por el gradiente menos $q\mathbf{A}$ en cada caso, de modo que la ecuación (21.2) se convierte en

$$-\frac{\hbar}{i} \frac{\partial \psi}{\partial t} = \hat{\mathcal{H}} \psi = \frac{1}{2m} \left(\frac{\hbar}{i} \nabla - q\mathbf{A} \right) \cdot \left(\frac{\hbar}{i} \nabla - q\mathbf{A} \right) \psi + q\phi\psi, \quad (21.3)$$

Ésta es la ecuación de Schrödinger para una partícula de carga q que se mueve en un campo electromagnético $\mathbf{A}$, ϕ (no relativista, sin espín).

Para demostrar que esto es verdad, lo ilustraré con un ejemplo simple en el que, en lugar de tener una situación continua, tendremos una línea de átomos según el eje x con un espaciamiento b y una amplitud $-K$ de que un electrón salte de un átomo a otro en ausencia del campo.[‡] Ahora bien, según la ecuación (21.1) si hay un potencial vectorial $A_x(x, t)$ en la dirección x , la amplitud de saltar variará respecto a su valor anterior en un factor $\exp[(iq/\hbar)A_x b]$, siendo el exponente $iq/\hbar$ por el potencial vectorial integrado desde un átomo al siguiente. Para simplificar escribiremos $(q/\hbar)A_x \equiv f(x)$, puesto que A_x dependerá en general de x . Si llamamos $C(x) \equiv C_n$ a la amplitud de encontrar el electrón en el átomo “ n ” ubicado en x , la derivada respecto al tiempo de dicha amplitud está dada por la ecuación siguiente:

$$\begin{aligned} -\frac{\hbar}{i} \frac{\partial}{\partial t} C(x) &= E_0 C(x) - K e^{-ibf(x+b/2)} C(x+b) \\ &\quad - K e^{+ibf(x-b/2)} C(x-b). \end{aligned} \quad (21.4)$$

Hay tres contribuciones. En primer lugar, hay una cierta energía E_0 si el electrón está ubicado en x . Como siempre, esto nos da el término $E_0 C(x)$. Luego tenemos el término $-K C(x+b)$ que es la amplitud de que el electrón haya dado un paso hacia atrás desde el átomo “ $n+1$ ”, ubicado en $x+b$. Sin embargo, al ocurrir eso en un potencial vectorial, la fase de la amplitud se debe correr de acuerdo con la regla de la ecuación (21.1). Si A_x no varía apreciablemente en un espaciamiento atómico, la integral se puede escribir simplemente como el valor de A_x en el punto medio por el espaciamiento b . De aquí que $(iq/\hbar)$ por la integral es justamente $ibf(x+b/2)$. Como el electrón está saltando hacia atrás, he mostrado este corrimiento de fase con el signo menos. Esto nos da la segunda contribución. Del mismo modo hay una cierta amplitud de que haya saltado

¹ Volumen II, sección 15-5.

[†] ¡No hay que confundirlo con nuestro uso anterior de ϕ para designar un estado!

[‡] K es la misma cantidad que llamamos A en el problema de una red lineal en ausencia de campo magnético. Véase el capítulo 13.

desde el otro lado, y aquí necesitamos el potencial vectorial a una distancia $(b/2)$ al otro lado de x por la distancia b . Esto da la tercera contribución. La suma da la ecuación para la amplitud de estar en x en un potencial vectorial.

Ahora bien, sabemos que si la función $C(x)$ es suficientemente suave (tomamos el límite de longitud de onda grande) y, si suponemos que los átomos están muy juntos, la ecuación (21.4) se aproximará al comportamiento de un electrón en el espacio libre. Por lo que el paso siguiente es desarrollar el lado derecho de (21.4) en potencias de b , suponiendo que b es muy pequeño. Por ejemplo, si b es cero el segundo miembro es justamente $(E_0 - 2K)C(x)$, por lo que en dicha aproximación la energía es $E_0 - 2K$. Después vienen los términos en b . Pero como las dos exponenciales tienen signos opuestos, solamente quedan las potencias pares de b . Así, pues, si hacen un desarrollo de Taylor de $C(x)$, de $f(x)$ y de las exponenciales y reúnen los términos en b^2 , obtienen

$$-\frac{\hbar}{i} \frac{\partial C(x)}{\partial t} = E_0 C(x) - 2K C(x) - Kb^2 \{C''(x) - 2if(x)C'(x) - if'(x)C(x) - f^2(x)C(x)\}. \quad (21.5)$$

(Las “primas” indican derivación respecto a x .)

Esta combinación horrible de cosas parece muy complicada. Pero matemáticamente es exactamente lo mismo que

$$-\frac{\hbar}{i} \frac{\partial C(x)}{\partial t} = (E_0 - 2K)C(x) - Kb^2 \left[\frac{\partial}{\partial x} - if(x) \right] \left[\frac{\partial}{\partial x} - if(x) \right] C(x). \quad (21.6)$$

Cuando el segundo corchete opera sobre $C(x)$ da $C'(x)$ más $if(x)C(x)$. El primer corchete operando en estos dos términos da el término C'' y términos en la primera derivada de $f(x)$ y en la primera derivada de $C(x)$. Recuerden que las soluciones en ausencia de campo magnético² representan una partícula con una masa efectiva m_{ef} dada por

$$Kb^2 = \frac{\hbar^2}{2m_{\text{ef}}}.$$

Si hacen entonces $E_0 = -2K$ e introducen $f(x) = (q/\hbar)A_x$ pueden comprobar fácilmente que la ecuación (21.6) es igual al primer miembro de la ecuación (21.3). (El origen del término de energía potencial se conoce bien, por lo que no me he molestado en introducirlo en esta discusión.) La proposición de la ecuación (21.1) de que el potencial vectorial cambia todas las amplitudes en el factor exponencial es igual a la regla de que el operador momentum $(\hbar/i)\nabla$ se reemplaza por

$$\frac{\hbar}{i} \nabla - q\mathbf{A},$$

como ven en la ecuación de Schrödinger de (21.3).

21-2 Ecuación de continuidad para probabilidades

Ahora, volvamos a un segundo punto. Una parte importante de la ecuación de Schrödinger para una sola partícula es la idea de que la probabilidad de encontrar la partícula en una posición está dada por el cuadrado del módulo de la función de onda. Es también característico de la mecánica cuántica que la probabilidad se conserve en un sentido local. Cuando disminuye la probabilidad de encontrar el electrón en algún lugar, mientras aumenta la probabilidad de que el electrón esté en cualquier otro lugar (conservándose la probabilidad total), algo debe estar ocurriendo en medio. Es decir, el electrón tiene una continuidad en el sentido de que si la probabilidad disminuye en un lugar y se forma en otro lugar, debe haber una cierta especie de flujo entre los dos. Por ejemplo, si colocan una pared en el camino, tendrá una influencia y las probabilidades no serán las mismas. Por lo tanto, la conservación de la probabilidad sola no

² Sección 13-3.

es el enunciado completo de la ley de conservación, del mismo modo que la conservación de la sola energía no es tan profunda e importante como la conservación *local* de la energía.³ Si alguna energía desaparece, entonces debe haber un flujo correspondiente de energía. Del mismo modo, quisiéramos encontrar una “corriente” de probabilidad tal que si hubiese cualquier cambio en la densidad de probabilidad (probabilidad de ser encontrado en una unidad de volumen), se le pudiese considerar como originado por el flujo entrante o saliente de una corriente. Esta corriente sería un vector que se podría interpretar así —la componente x sería la probabilidad total por segundo y por unidad de área de que una partícula atravesase en la dirección x un plano paralelo al plano y - z —. Cuando pasa hacia $+x$ se considera un flujo positivo, negativo cuando pasa en dirección opuesta. ¿Hay tal corriente? Bueno, saben que la densidad de probabilidad $P(\mathbf{r}, t)$ está dada en términos de la función de onda por

$$P(\mathbf{r}, t) = \psi^*(\mathbf{r}, t)\psi(\mathbf{r}, t). \quad (21.7)$$

Mi pregunta es: ¿hay una corriente $\mathbf{J}$ tal que

$$\frac{\partial P}{\partial t} = -\nabla \cdot \mathbf{J}? \quad (21.8)$$

Si derivó la ecuación (21.7) respecto al tiempo, obtengo dos términos:

$$\frac{\partial P}{\partial t} = \psi^* \frac{\partial \psi}{\partial t} + \psi \frac{\partial \psi^*}{\partial t}. \quad (21.9)$$

Ahora usen la ecuación de Schrödinger —ecuación (21.3)— para $\partial \psi / \partial t$, tomen el complejo conjugado para obtener $\partial \psi^* / \partial t$ —cada i invierte su signo—. Obtienen

$$\begin{aligned} \frac{\partial P}{\partial t} = & -\frac{i}{\hbar} \left[\psi^* \frac{1}{2m} \left(\frac{\hbar}{i} \nabla - q\mathbf{A} \right) \cdot \left(\frac{\hbar}{i} \nabla - q\mathbf{A} \right) \psi + q\phi \psi^* \psi \right. \\ & \left. - \psi \frac{1}{2m} \left(-\frac{\hbar}{i} \nabla - q\mathbf{A} \right) \cdot \left(-\frac{\hbar}{i} \nabla - q\mathbf{A} \right) \psi^* - q\phi \psi \psi^* \right]. \end{aligned} \quad (21.10)$$

Los términos potenciales y muchas otras cosas se cancelan. Y resulta que lo que queda se puede realmente escribir como una divergencia perfecta. La ecuación entera es equivalente a

$$\frac{\partial P}{\partial t} = -\nabla \cdot \left\{ \frac{1}{2m} \psi^* \left(\frac{\hbar}{i} \nabla - q\mathbf{A} \right) \psi + \frac{1}{2m} \psi \left(-\frac{\hbar}{i} \nabla - q\mathbf{A} \right) \psi^* \right\}. \quad (21.11)$$

Realmente no es tan complicada como parece. Es una combinación simétrica de ψ^* por una cierta operación sobre ψ , más ψ por la operación compleja conjugada sobre ψ^* . Es una cantidad más su propia compleja conjugada, por lo que el resultado es real —como debe ser—. La operación se puede recordar así: es solamente el operador momentum $\hat{\mathbf{P}}$ menos $q\mathbf{A}$. Podría escribir la corriente de la ecuación (21.8) en la forma

$$\mathbf{J} = \frac{1}{2} \left\{ \psi^* \left[\frac{\hat{\mathbf{P}} - q\mathbf{A}}{m} \right] \psi + \psi \left[\frac{\hat{\mathbf{P}} - q\mathbf{A}}{m} \right]^* \psi^* \right\}. \quad (21.12)$$

Hay entonces una corriente $\mathbf{J}$ que completa la ecuación (21.8).

La ecuación (21.11) nos dice que la probabilidad se conserva localmente. Si una partícula desaparece de una región, no puede aparecer en otra a menos que algo fluya entre ellas. Imaginen que la primera está rodeada por una superficie cerrada lo bastante alejada para que haya probabilidad cero de encontrar el electrón en la superficie. La probabilidad total de encontrar el electrón en algún lugar dentro de la superficie es la integral de volumen de P . Pero según el teorema de Gauss la integral de volumen de la divergencia de $\mathbf{J}$ es igual a la integral de superficie de $\mathbf{J}$. Si ψ es cero en la superficie, la ecuación (21.12) expresa que $\mathbf{J}$ es cero, por lo que la probabilidad total de encontrar la partícula en el interior no puede variar. Sólo puede gotear algo de probabilidad hacia afuera si alguna probabilidad se acerca a la superficie. Podemos decir que sólo sale si atraviesa la superficie —y esto es conservación local—.

³ Volumen II, sección 27-1.

21-3 Dos clases de momentum

La ecuación para la corriente es bastante interesante y causa a veces cierta dificultad. Pueden pensar que la corriente sería algo así como la densidad de partículas por la velocidad. La densidad sería algo así como $\psi\psi^*$, lo cual está bien. Y cada término de la ecuación (21.12) se parece a la forma típica para el valor medio del operador

$$\frac{\hat{\mathcal{P}} - q\mathbf{A}}{m}, \quad (21.13)$$

por lo que quizá podríamos considerarlo como la velocidad de flujo. Al parecer tenemos dos sugerencias para relacionar la velocidad con el momentum, ya que también podríamos considerar que el momentum dividido entre la masa, $\hat{\mathcal{P}}/m$, debería ser una velocidad. Las dos posibilidades difieren en el potencial vectorial.

Resulta que estas dos posibilidades también se descubrieron en la física clásica cuando se encontró que el momentum se podía definir de dos formas.⁴ Uno de ellos se llama “momentum cinemático”, pero para ser absolutamente claro lo llamaré en este capítulo “momentum mv ”. Éste es el momentum que se obtiene multiplicando la masa por la velocidad. El otro, que se denomina “momentum dinámico” y lo llamaré “momentum p ”, es un momentum más abstracto, más matemático. Las dos posibilidades son

$$\text{momentum } mv = m\mathbf{v}, \quad (21.14)$$

$$\text{momentum } p = m\mathbf{v} + q\mathbf{A}. \quad (21.15)$$

Ocurre que en la mecánica cuántica con campos magnéticos presentes el momentum p es el que se relaciona con el operador gradiente $\hat{\mathcal{P}}$, por lo que (21.13) es el operador de una velocidad.

Quisiera hacer una breve digresión para mostrarles de qué se trata —por qué debe haber algo como la ecuación (21.15) en la mecánica cuántica—. La función de onda varía en el tiempo, de acuerdo con la ecuación de Schrödinger (21.3). Si de repente yo cambiase el potencial vectorial, la función de onda no variaría instantáneamente; solamente variaría su rapidez de variación. Consideren ahora lo que sucedería en la circunstancia siguiente. Supongan que tengo un largo solenoide en el que puedo producir un flujo de campo magnético (campo $\mathbf{B}$), como muestra la figura 21-2, y que hay una partícula cargada quieta en sus proximidades. Supongan que se crea este flujo casi instantáneamente llevándolo desde cero a algún valor. Comienzo con un potencial vectorial nulo y luego tengo un potencial vectorial. Esto significa que produzco de repente un potencial vectorial circunferencial $\mathbf{A}$. Recordarán que la integral de línea de $\mathbf{A}$ según un lazo es igual al flujo de $\mathbf{B}$ a través del lazo.⁵ Ahora bien, ¿qué sucederá si de repente introduzco un potencial vectorial? Según la ecuación cuántica, el cambio brusco de $\mathbf{A}$ no produce un cambio brusco en ψ ; la función de onda es aún la misma. Por lo que el gradiente tampoco varía.

Pero recuerden lo que sucede eléctricamente cuando introduzco de repente un flujo. Durante el pequeño intervalo de tiempo en que el flujo está aumentando, se genera un campo eléctrico cuya integral de línea es la derivada del flujo respecto al tiempo:

$$\mathbf{E} = -\frac{\partial \mathbf{A}}{\partial t}. \quad (21.16)$$

Este campo eléctrico es enorme si el flujo varía rápidamente y produce una fuerza sobre la partícula. La fuerza es la carga por el campo eléctrico y por ello durante la introducción del flujo la partícula alcanza un impulso total (es decir, una variación de $m\mathbf{v}$) igual a $-q\mathbf{A}$. Dicho de otro modo, si de repente introducen un potencial vectorial afectando una carga, esta carga toma inmediatamente un momentum mv igual a $-q\mathbf{A}$. Pero hay algo que no varía inmediatamente y eso

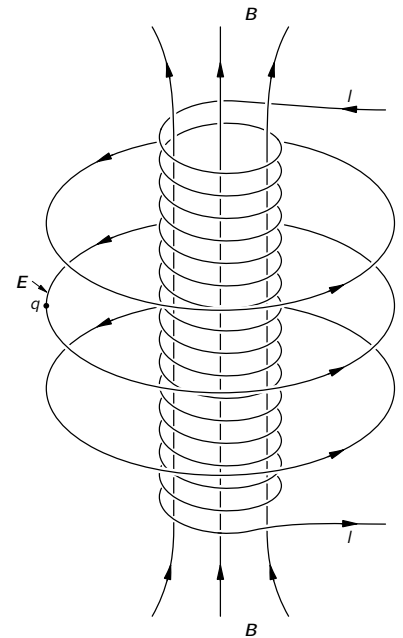


Figura 21-2. El campo eléctrico fuera de un solenoide con una corriente creciente.

⁴ Véase, por ejemplo, J. D. Jackson, *Classical Electrodynamics*, John Wiley and Sons, Inc., Nueva York, 1962, p. 408.

⁵ Volumen II, capítulo 14, sección 14-1.

es la diferencia entre $m\mathbf{v}$ y $-q\mathbf{A}$. Por lo tanto, la suma $\mathbf{p} = m\mathbf{v} + q\mathbf{A}$ es algo que no varía cuando efectúan un cambio brusco en el potencial vectorial. Esta cantidad $\mathbf{p}$ es la que hemos denominado momentum p y tiene importancia en la mecánica clásica en la teoría de la dinámica, pero tiene también un significado directo en la mecánica cuántica. Depende del carácter de la función de onda y es la que se identifica con el operador

$$\hat{\mathbf{p}} = \frac{\hbar}{i} \nabla.$$

21-4 El significado de la función de onda

Cuando Schrödinger descubrió su ecuación, descubrió la ley de conservación (21.8) como una consecuencia de la misma. Pero incorrectamente imaginó que P era la densidad de carga eléctrica del electrón y que $\mathbf{J}$ era la densidad de corriente eléctrica, por lo que pensó que los electrones interactuaban con el campo electromagnético mediante estas cargas y corrientes. Cuando resolvió su ecuación para el átomo de hidrógeno y calculó ψ , no estaba calculando la probabilidad de nada —en aquel tiempo no había amplitudes— y la interpretación fue completamente diferente. El núcleo atómico estaba quieto, pero había corrientes circulando a su alrededor; las cargas P y las corrientes $\mathbf{J}$ generarían campos electromagnéticos y todo esto radiaría luz. Pronto encontró al resolver algunos problemas que no funcionaba muy bien. Fue aquí cuando Born contribuyó esencialmente a nuestras ideas respecto a la mecánica cuántica. Fue Born quien interpretó correctamente (hasta donde sabemos) la ψ de la ecuación de Schrödinger en términos de una amplitud de probabilidad —esa idea tan difícil de que el cuadrado de la amplitud no es la densidad de carga, sino solamente la probabilidad por unidad de volumen de encontrar un electrón allí y cuando encuentran un electrón en algún lugar la carga total está allí—. Todo esto constituye un aporte de Born.

La función de onda $\psi(\mathbf{r})$ de un electrón en un átomo no describe, pues, un electrón diseminado con una densidad de carga uniforme. El electrón o está aquí, o está allí o en cualquier otro lugar, pero dondequiera que esté es una carga puntual. Por otra parte, consideren la situación en que hay un gran número de partículas exactamente en el mismo estado, un gran número de ellas son exactamente la misma función de onda. ¿Qué pasa entonces? Una de ellas está aquí y otra está allá, y la probabilidad de encontrar cualquiera de ellas en un lugar dado es proporcional a $\psi\psi^*$. Pero como hay tantas partículas, si miro cualquier volumen $dx\,dy\,dz$, encontraré generalmente un número próximo a $\psi\psi^* dx\,dy\,dz$. Por consiguiente, en una situación en que ψ es la función de onda para cada una de las partículas de entre un gran número de ellas que están todas en el mismo estado, $\psi\psi^*$ se puede interpretar como la densidad de partículas. Si, en estas circunstancias, cada partícula lleva la misma carga q , podemos de hecho avanzar más e interpretar $\psi^*\psi$ como la densidad de *electricidad*. Normalmente, $\psi\psi^*$ tiene las dimensiones de una densidad de probabilidad y ψ se debería multiplicar por q para que dé las dimensiones de una densidad de carga. Para nuestros fines presentes podemos introducir este factor constante en ψ y tomar $\psi\psi^*$ como la densidad de carga eléctrica. De este modo, $\mathbf{J}$ (la probabilidad de corriente que he calculado) pasa directamente a ser la densidad de corriente eléctrica.

Por lo tanto, en el caso de tener muchísimas partículas exactamente en el mismo estado, es posible una interpretación física nueva de las funciones de onda. La densidad de carga y la corriente eléctrica se pueden calcular directamente a partir de las funciones de onda y éstas toman un significado físico que se extiende a situaciones macroscópicas clásicas.

Algo parecido puede suceder con partículas neutras. Cuando tenemos la función de onda de un solo fotón, es la amplitud de encontrar el fotón en algún lugar. Aunque nunca la hemos escrito, hay una ecuación para la función de onda del fotón análoga a la ecuación de Schrödinger para el electrón. La ecuación del fotón es exactamente lo mismo que las ecuaciones de Maxwell para el campo

electromagnético y la función de onda es lo mismo que el potencial vectorial $\mathbf{A}$. La función de onda resulta ser exactamente el potencial vectorial. La física cuántica es igual a la física clásica porque los fotones son bosones que no interactúan y muchos de ellos pueden encontrarse en el mismo estado —como saben, les *gusta* estar en el mismo estado—. En el momento que tengan miles de millones en el mismo estado (es decir, en la misma onda electromagnética) pueden medir la función de onda, o sea el potencial vectorial, directamente. Como es natural, históricamente se hizo del otro modo. Las primeras observaciones ocurrieron en situaciones con muchos fotones en el mismo estado y por ello pudimos descubrir la ecuación correcta para un solo fotón, observando directamente con nuestras manos la naturaleza de la función de onda a un nivel macroscópico.

Ahora bien, la dificultad con el electrón está en que no pueden colocar más de un electrón en el mismo estado. Por consiguiente, se creyó por mucho tiempo que la función de onda de la ecuación de Schrödinger nunca tendría una representación macroscópica análoga a la representación macroscópica de la amplitud para fotones. Por otra parte, se ha visto ahora que los fenómenos de superconductividad nos muestran exactamente esta situación.

21-5 Superconductividad

Como saben, muchísimos metales pasan a ser superconductores por debajo de una cierta temperatura⁶ —la temperatura es diferente para distintos metales—. Cuando reducen suficientemente la temperatura, los metales conducen la electricidad sin ninguna resistencia. Este fenómeno se ha observado en un gran número de metales, pero no en todos, y la teoría de este fenómeno ha originado muchísima dificultad. Consumió mucho tiempo entender qué era lo que pasaba dentro de los superconductores y aquí sólo describiré lo suficiente para nuestros propósitos. Resulta que, debido a las interacciones de los electrones con las vibraciones de los átomos de la red, hay una pequeña *atracción* efectiva resultante entre los electrones. El resultado es que los electrones forman pares ligados, si se me permite hablar muy cualitativa y fríamente.

Ahora bien, ustedes saben que un electrón es un fermión. Pero un par ligado actuaría como un bosón, puesto que si intercambio los electrones de un par, cambio los signos de la función de onda dos veces y esto significa que no cambio nada. Un par *es* un bosón.

La energía de apareamiento —es decir, la atracción resultante— es muy, muy débil. Se necesita sólo una pequeñísima temperatura para separar los electrones debido a la agitación térmica y convertirlos en electrones “normales”. Pero cuando se disminuye la temperatura lo suficiente como para que tengan que hacer lo posible para ir al estado absolutamente más bajo, entonces se reúnen en pares.

No quiero que piensen que los pares están realmente tan unidos que forman una partícula puntual. En realidad, una de las grandes dificultades para la comprensión de este fenómeno originalmente fue la de que no es así como están. Los dos electrones que forman el par están realmente separados a una gran distancia; y la distancia media entre pares es relativamente más pequeña que el tamaño de un único par. Varios pares ocupan el mismo espacio al mismo tiempo. La razón de por qué los electrones forman pares y una estimación de la energía necesaria para formarlos han sido triunfos recientes. Bardeen, Cooper y Schrieffer⁷ explicaron por primera vez este punto fundamental de la teoría de la superconductividad, pero éste no es el tema de este seminario. Sin embargo, aceptaremos la idea de que los electrones trabajan realmente en pares, de una forma u otra, de que podemos considerar que estos pares se comportan más o menos como partículas y de que, por consiguiente, podemos hablar de función de onda de un “par”.

⁶ Descubierto por primera vez por Kamerlingh-Onnes en 1911; H. Kamerlingh-Onnes, *Comm. Phys. Lab., Univ. Leyden*, núms. 119, 120, 122 (1911). Se puede encontrar una buena discusión actualizada del tema en E. A. Lynton, *Superconductivity*, John Wiley and Sons, Inc., Nueva York, 1962.

⁷ J. Bardeen, L. N. Cooper y J. R. Schrieffer, *Phys. Rev.* **108**, 1175 (1957).

Ahora bien, la ecuación de Schrödinger para el par será más o menos como la ecuación (21.3). Habrá una diferencia de que la carga q será el doble de la carga de un electrón. Asimismo, no conocemos la inercia —o masa efectiva— del par en la red cristalina, por lo que no sabemos qué número colocar para m . Tampoco deberíamos pensar que si vamos a frecuencias muy altas (o pequeñas longitudes de onda) ésta es exactamente la forma correcta, ya que la energía cinética que corresponde a funciones de onda que varían muy rápidamente puede ser tan grande como para romper los pares. A temperaturas finitas hay siempre algunos pares que se rompen de acuerdo a la teoría de Boltzmann. La probabilidad de que se rompa un par es proporcional a $\exp(-E_{\text{par}}/\kappa T)$. A los electrones que no forman pares se les llama electrones “normales” y se moverán por el cristal en la forma acostumbrada. Sin embargo, consideraré solamente el caso de temperatura esencialmente cero —o, de todas maneras, despreciaré las complicaciones producidas por los electrones que no forman pares—.

Como los pares de electrones son bosones, cuando hay una gran cantidad de ellos en un estado dado hay una amplitud especialmente grande de que otros pares vayan al mismo estado. Por lo tanto, todos los pares se verán confinados a la energía más baja en *exactamente el mismo estado* —no será fácil obtener uno de ellos en otro estado—. Hay más amplitud de ir al mismo estado que a un estado desocupado en el factor famoso $\sqrt{n}$, donde $n-1$ es el número de ocupación del estado más bajo. Así, esperaríamos que todos los pares se trasladasen al mismo estado.

¿Cuál será el aspecto de nuestra teoría? Llamaré ψ a la función de onda de un par en el estado más bajo de energía. Sin embargo, como $\psi\psi^*$ va a ser proporcional a la densidad de carga ρ , puedo escribir también ψ como la raíz cuadrada de la densidad de carga por algún factor de fase:

$$\psi(\mathbf{r}) = \rho^{1/2}(\mathbf{r})e^{i\theta(\mathbf{r})}, \quad (21.17)$$

donde ρ y θ son funciones reales de $\mathbf{r}$. (Cualquier función compleja se puede, naturalmente, escribir de este modo.) Es claro lo que queremos dar a entender cuando hablamos de la densidad de carga, no obstante ¿cuál es el significado físico de la fase θ de la función de onda? Bien, veamos que pasa si sustituimos $\psi(\mathbf{r})$ en la ecuación (21.12) y expresamos la densidad de corriente en función de estas dos nuevas variables ρ y θ . Es sólo un cambio de variables y evitaré toda el álgebra, pero resulta

$$\mathbf{J} = \frac{\hbar}{m} \left(\nabla\theta - \frac{q}{\hbar} \mathbf{A} \right) \rho. \quad (21.18)$$

Puesto que tanto la densidad de corriente como la densidad de carga tienen un significado físico directo para el gas electrónico superconductor, tanto ρ como θ son cosas reales. La fase es exactamente tan observable como ρ ; es una parte de la densidad de corriente $\mathbf{J}$. La fase *absoluta* no es observable, pero si se conoce en todo punto el gradiente de la fase, se conoce la fase excepto por una constante. Pueden definir la fase en un punto y estará determinada en cualquier punto.

Incidentalmente, la ecuación para la corriente se puede analizar un poco más elegantemente, si piensan que la densidad de corriente $\mathbf{J}$ es *de hecho* la densidad de carga por la velocidad de desplazamiento del fluido electrónico, o sea $\rho\mathbf{v}$. La ecuación (21.18) es entonces equivalente a

$$m\mathbf{v} = \hbar \nabla\theta - q\mathbf{A}. \quad (21.19)$$

Observen que hay dos términos en el momentum $m\mathbf{v}$; uno es una contribución debida al potencial vectorial y el otro una contribución debida al comportamiento de la función de onda. En otras palabras, la cantidad $\hbar\nabla\theta$ es simplemente lo que hemos llamado momentum p .

21-6 El efecto Meissner

Ahora podemos describir algunos fenómenos de superconductividad. En primer lugar, no hay resistencia eléctrica. Y no la hay porque todos los electrones

están colectivamente en el mismo estado. En el flujo ordinario de corriente pueden sacar un electrón o el otro del flujo regular, deteriorando así gradualmente el momentum general. Pero aquí, separar un electrón de lo que están haciendo todos los demás es muy difícil debido a la tendencia de todos los bosones a ir al mismo estado. Una vez que empieza una corriente, no se destruye jamás.

También es fácil entender que si tienen un trozo de metal en el estado superconductor e introducen un campo magnético que no sea muy intenso (no entraremos en detalles de cuán intenso), el campo magnético no puede penetrar en el metal. Si a medida que se introduce el campo magnético algo de él se crease dentro del metal, habría una velocidad de variación de flujo que produciría un campo eléctrico y éste generaría inmediatamente una corriente que, según la ley de Lenz, se opondría al flujo. Como todos los electrones se mueven a la vez, un campo eléctrico infinitesimal generará corriente suficiente para anular completamente cualquier campo magnético aplicado. Por consiguiente, si aplican el campo después que han enfriado un metal hasta el estado superconductor, el campo no podrá penetrar.

Más interesante aún es un fenómeno relacionado descubierto experimentalmente por Meissner.⁸ Si tienen un trozo del metal a una temperatura alta (y, por consiguiente, es un conductor normal) y establecen un campo magnético a través de él y luego bajan la temperatura por debajo de la temperatura crítica (cuando el metal pasa a ser superconductor), *el campo es expulsado*. En otras palabras, aquél origina su propia corriente —y en la cantidad justa para expulsar el campo—.

Nos daremos cuenta de ello por las ecuaciones y querría explicarles cómo. Supongan que tomamos un trozo de metal superconductor que forma un todo. Entonces, en una situación estable de cualquier clase, la divergencia de la corriente debe ser cero porque no tiene donde ir. Es conveniente escoger que la divergencia de $\mathbf{A}$ sea igual a cero (debería explicar por qué si escogemos esta convención no hay pérdida de generalidad, pero no lo haré). Si tomamos la divergencia de la ecuación (21.18) nos da que el laplaciano de θ es igual a cero. Pero... un momento. ¿Qué hay de la variación de ρ ? Olvidé mencionar un punto importante. Hay un fondo de carga positiva en este metal debido a los iones atómicos de la red. Si la densidad ρ de carga es uniforme, no hay ni carga neta ni campo eléctrico. Si hubiese cualquier acumulación de electrones en una región, la carga no se neutralizaría y habría una repulsión impresionante tratando de separar los electrones.[†] Así, en circunstancias comunes la densidad de carga de los electrones en el superconductor es casi perfectamente uniforme —puedo tomar ρ como constante—. Ahora bien, el único modo de que $\nabla^2\theta$ pueda ser cero en todos los puntos interiores del trozo de metal es si θ es constante. Y esto significa que el momentum p no contribuye a $\mathbf{J}$. La ecuación (21.18) nos dice, desde luego, que la corriente es proporcional a ρ por $\mathbf{A}$. Por lo tanto, en un trozo de material superconductor la corriente es necesariamente proporcional al potencial vectorial:

$$\mathbf{J} = -\rho \frac{q}{m} \mathbf{A}. \quad (21.20)$$

Como ρ y q tienen el mismo signo (negativo) y como ρ es una constante, puedo poner $-\rho q/m = -(\text{cierta constante positiva})$; entonces

$$\mathbf{J} = -(\text{cierta constante positiva})\mathbf{A}. \quad (21.21)$$

London y London⁹ propusieron esta ecuación originalmente para explicar las observaciones experimentales de superconductividad —mucho antes de que se entendiera el origen cuántico del efecto—.

⁸ W. Meissner y R. Ochsenfeld, *Naturwiss.* **21**, 787 (1933).

[†] En realidad si el campo eléctrico fuese demasiado intenso se destruirían algunos pares y los electrones “normales” creados se moverían para ayudar a neutralizar cualquier exceso de carga positiva. Pero aun así, se necesita energía para obtener estos electrones normales, de modo que lo importante es que se favorece muchísimo por razones energéticas una densidad ρ casi uniforme.

⁹ F. London y H. London, *Proc. Roy. Soc. (London)* **A149**, 71 (1935); *Physica* **2**, 341 (1935).

Ahora podemos introducir la ecuación (21.20) en las ecuaciones del electromagnetismo y calcular los campos. El potencial vectorial se relaciona con la densidad de corriente media de la siguiente forma

$$\nabla^2 \mathbf{A} = -\frac{1}{\epsilon_0 c^2} \mathbf{J}. \quad (21.22)$$

Si uso la ecuación (21.21) para $\mathbf{J}$, tengo

$$\nabla^2 \mathbf{A} = \lambda^2 \mathbf{A}, \quad (21.23)$$

donde λ^2 es simplemente una constante nueva;

$$\lambda^2 = \rho \frac{q}{\epsilon_0 m c^2}. \quad (21.24)$$

Ahora podemos tratar de resolver esta ecuación para encontrar $\mathbf{A}$ y ver en detalle lo que sucede. Por ejemplo, la ecuación (21.23) en una dimensión tiene soluciones exponenciales de la forma $e^{-\lambda x}$ y $e^{+\lambda x}$. Estas soluciones indican que el potencial vectorial debe *disminuir* exponencialmente cuando se entra desde la superficie al material. (No puede aumentar porque habría una explosión.) Si el trozo de metal es muy grande comparado con $1/\lambda$, el campo penetra solamente un espesor muy delgado —un espesor de alrededor de $1/\lambda$ —. Todo el resto del interior está libre de campo como representa la figura 21-3. Ésta es la explicación del efecto Meissner.

¿Qué tan grande es la distancia λ ? Bien, recuerden que r_0 , “radio electromagnético” del electrón (2.8×10^{-13} cm), está dado por

$$m c^2 = \frac{q_e^2}{4\pi\epsilon_0 r_0}.$$

Recuerden también que q de la ecuación (21.24) es el doble de la carga de un electrón, por lo que

$$\frac{q^2}{\epsilon_0 m c^2} = \frac{8\pi r_0}{q_e}.$$

Escribiendo ρ como $q_e N$, donde N es el número de electrones por centímetro cúbico, tenemos

$$\lambda^2 = 8\pi N r_0. \quad (21.25)$$

Para un metal como el plomo hay alrededor de 3×10^{22} átomos por cm^3 , por lo que si cada uno contribuye solamente con un electrón de conducción, $1/\lambda$ sería alrededor de 2×10^{-6} cm. Esto les da el orden de magnitud.

21-7 Cuantificación del flujo

La ecuación de London (21.21) se propuso para tener en cuenta los efectos de superconductividad observados, incluyendo el efecto Meissner. Sin embargo, en tiempos recientes se han dado predicciones aún más espectaculares. Una dada por London era tan peculiar que nadie le prestó mucha atención hasta hace poco. La discutiremos ahora. Pero esta vez, en lugar de tomar un trozo, tomamos un *anillo* cuyo espesor es grande comparado con $1/\lambda$ y tratamos de ver lo que sucedería si tuviéramos un campo magnético atravesando el anillo, después lo enfriáramos hasta el estado superconductor y luego quitáramos la fuente original de $\mathbf{B}$. La figura 21-4 muestra la sucesión de los hechos. En el estado normal habrá un campo en el interior del anillo como muestra la parte (a) de la figura. Cuando el anillo se hace superconductor, el campo es forzado fuera del *material* (como acabamos de ver). Habrá entonces algún flujo a través del agujero del anillo, como muestra la parte (b). Si quitamos ahora el campo externo, las líneas de fuerza que atraviesan el agujero quedan “atrapadas”, como muestra la parte (c). El flujo Φ a través del centro no puede disminuir ya que $\partial\Phi/\partial t$ debe ser igual a la integral de línea de $\mathbf{E}$ alrededor del anillo que es cero en un superconductor. A medida que removemos el campo externo se origina una supercorriente que

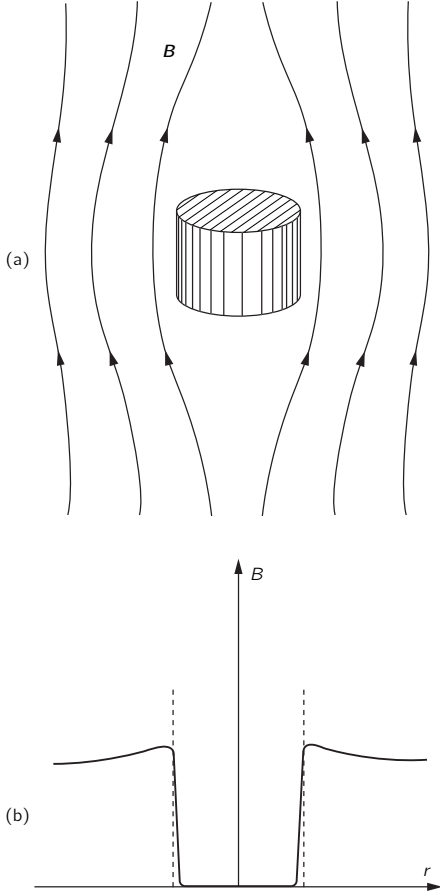


Figura 21-3. (a) Un cilindro superconductor en un campo magnético; (b) el campo magnético B como función de r .

fluye alrededor del anillo para conservar constante el flujo a través del anillo (es la vieja idea de corrientes parásitas, sólo que con resistencia cero). Sin embargo, estas corrientes fluirán cerca de la superficie (hasta una profundidad $1/\lambda$), como se puede demostrar siguiendo el mismo análisis que utilicé para un bloque sólido. Estas corrientes pueden mantener el campo magnético fuera del anillo y producir igualmente el campo magnético atrapado permanentemente.

No obstante, ahora hay una diferencia esencial y nuestras ecuaciones predicen un efecto sorprendente. El razonamiento que hice anteriormente de que θ debe ser constante en un bloque sólido *no sirve para un anillo*, como pueden ver por los razonamientos siguientes.

En el interior del anillo la densidad de corriente $\mathbf{J}$ es cero; por lo que la ecuación (21.18) nos da

$$\hbar \nabla \theta = q \mathbf{A}. \quad (21.26)$$

Consideren ahora lo que obtenemos si tomamos la integral de línea de $\mathbf{A}$ en una curva Γ , alrededor del anillo cerca del centro de su sección y nunca cerca de la superficie, como muestra la figura 21-5. Según la ecuación (21.26),

$$\hbar \oint \nabla \theta \cdot d\mathbf{s} = q \oint \mathbf{A} \cdot d\mathbf{s}. \quad (21.27)$$

Ahora bien, saben que la integral de línea de $\mathbf{A}$ siguiendo cualquier camino cerrado es igual al flujo de $\mathbf{B}$ a través de dicha espira

$$\oint \mathbf{A} \cdot d\mathbf{s} = \Phi.$$

La ecuación (21.27) se convierte entonces en

$$\oint \nabla \theta \cdot d\mathbf{s} = \frac{q}{\hbar} \Phi. \quad (21.28)$$

La integral de línea de un gradiente desde un punto hasta otro (digamos del punto 1 al punto 2) es la diferencia de los valores de la función en los dos puntos. Es decir,

$$\int_1^2 \nabla \theta \cdot d\mathbf{s} = \theta_2 - \theta_1.$$

Si escogemos los dos puntos extremos 1 y 2 coincidentes de modo que formen un lazo, podrían pensar a primera vista que θ_2 sería igual a θ_1 y de ese modo la integral de la ecuación (21.28) sería cero. Esto sería válido para un lazo en un pedazo simplemente conexo del superconductor, pero no es necesariamente válido para un pedazo en forma de anillo. El único requisito físico que podemos hacer es que *puede haber solamente un valor de la función de onda para cada punto*. Haga lo que haga θ , cuando dan una vuelta al anillo, al llegar al punto de partida la θ que obtengan debe dar el mismo valor para la función de onda

$$\psi = \sqrt{\rho} e^{i\theta}.$$

Esto sucederá si θ varía en $2\pi n$, siendo n cualquier entero. Así, si damos una vuelta completa al anillo, el primer miembro de la ecuación (21.27) debe ser $\hbar \cdot 2\pi n$. Usando la ecuación (21.28), obtengo que

$$2\pi n \hbar = q \Phi. \quad (21.29)$$

¡El flujo atrapado debe ser siempre un entero por $2\pi\hbar/q$! Si considerasen el anillo como un objeto clásico con una conductividad perfecta idealmente (es decir, infinita), pensarían que cualquiera que fuese el flujo que hubiese inicialmente, éste se conservaría —no se podría atrapar ninguna cantidad de flujo—. Pero la teoría cuántica de la superconductividad dice que el flujo puede ser cero, o $2\pi\hbar/q$, o $4\pi\hbar/q$, o $6\pi\hbar/q$, etc., pero no puede ser ningún valor intermedio. Debe ser un múltiplo de una unidad cuántica básica.

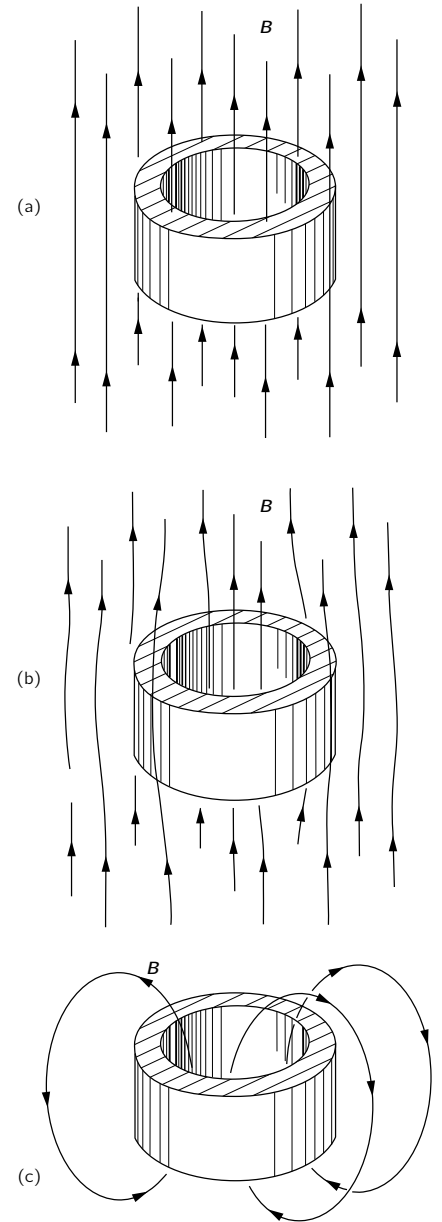


Figura 21-4. Un anillo en un campo magnético: (a) en el estado normal; (b) en el estado superconductor; (c) después de retirar el campo externo.

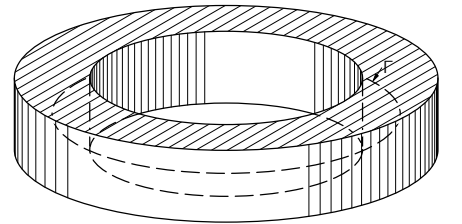


Figura 21-5. La curva Γ dentro de un anillo superconductor.

London¹⁰ predijo que el flujo atrapado por un anillo superconductor estaría cuantizado y dijo que los valores posibles de flujo estarían dados por la ecuación (21.29) con q igual a la carga electrónica. Según London, la unidad básica de flujo debería ser $2\pi\hbar/q_e$, que es alrededor de 4×10^{-7} gauss $\cdot$ cm². Para tener una idea gráfica de este flujo, consideren un cilindro delgado de una décima de milímetro de diámetro; cuando contiene esta cantidad de flujo, el campo magnético en su interior es aproximadamente uno por ciento del campo magnético terrestre. Debería ser posible observar dicho flujo mediante una medida magnética sensible.

Deaver y Fairbank¹¹ en la Universidad de Stanford, buscaron y encontraron este flujo cuantizado en 1961 y casi al mismo tiempo lo hicieron Doll y Näbauer¹² en Alemania.

En el experimento de Deaver y Fairbank se utilizó como superconductor un cilindro pequeño, construido galvanizando una delgada capa de estaño sobre un alambre de cobre de un centímetro de longitud, núm. 56 (1.3×10^{-3} cm de diámetro). El estaño se hace superconductor a una temperatura inferior a 3.8 K, mientras que a dicha temperatura el cobre sigue siendo un metal normal. Se colocó el alambre en un pequeño campo magnético controlado y se redujo la temperatura hasta que el estaño se convirtió en superconductor. Entonces se quitó la fuente externa que producía el campo. Sería de esperar que esto generase una corriente dada por la ley de Lenz de modo que no variase el flujo en el interior. El pequeño cilindro debería tener ahora un momento magnético proporcional al flujo en su interior. El momento magnético se midió moviendo hacia arriba y hacia abajo (como la aguja de una máquina de coser, pero a una media de 100 ciclos por segundo) el alambre entre un par de pequeñas bobinas en los extremos del cilindro de estaño. El voltaje inducido en las bobinas era una medida del momento magnético.

Cuando Deaver y Fairbank hicieron el experimento, encontraron que el flujo estaba cuantizado, *pero la unidad básica fue solamente la mitad del valor predicho por London*. Doll y Näbauer obtuvieron el mismo resultado. Al principio esto resultó misterioso,† aunque nadie entendió jamás por qué, pero nosotros entendemos ahora por qué debía ser así. Según la teoría de la superconductividad de Bardeen, Cooper y Schrieffer, la q que aparece en la ecuación (21.29) es la carga de un par de electrones y, por lo tanto, es igual a $2q_e$. La unidad básica de flujo es:

$$\Phi_0 = \frac{\pi\hbar}{q_e} \approx 2 \times 10^{-7} \text{ gauss} \cdot \text{cm}^2 \quad (21.30)$$

o sea, la mitad de lo que London predijo. Todo encaja ahora y las medidas muestran a escala macroscópica la existencia del efecto puramente cuántico predicho.

21-8 Dinámica de la superconductividad

El efecto Meissner y la cuantización del flujo son dos confirmaciones de nuestras ideas generales. Sólo por razones de totalidad nos gustaría mostrarles cuáles serían, desde este punto de vista, las ecuaciones completas de un fluido superconductor —y es muy interesante—. Hasta aquí he escrito solamente la expresión de ψ en ecuaciones de densidad de carga y corriente. Si la coloco en la ecuación completa de Schrödinger, obtengo ecuaciones para ρ y θ . Sería interesante ver qué resulta, ya que aquí tenemos un “fluido” de pares de electrones con una densidad de carga ρ y un θ misterioso —¡podemos tratar de ver qué clase de ecuaciones obtenemos para ese “fluido”!—. Así, sustituimos la función de onda de la ecuación (21.17) en la ecuación de Schrödinger (21.3), teniendo presente que ρ y θ son funciones reales de x , y , z y t . Si separamos las partes real

¹⁰ F. London, *Superfluids*; vol. I, John Wiley y Sons, Inc., Nueva York, 1950, p. 152.

¹¹ B. S. Deaver, Jr., y W. M. Fairbank, *Phys. Rev. Letters* **7**, 43 (1961).

¹² R. Doll y M. Näbauer, *Phys. Rev. Letters* **7**, 51 (1961).

† Onsager sugirió una vez que esto podría suceder (véase Deaver y Fairbank, nota 11)

e imaginaria obtenemos dos ecuaciones. Con el fin de escribirlas en una forma más corta —de acuerdo con la ecuación (21.19)— ponemos

$$\frac{\hbar}{m} \nabla \theta - \frac{q}{m} \mathbf{A} = \mathbf{v}. \quad (21.31)$$

Una de las ecuaciones que obtengo es entonces

$$\frac{\partial \rho}{\partial t} = -\nabla \cdot \rho \mathbf{v}. \quad (21.32)$$

Como $\rho \mathbf{v}$ es $\mathbf{J}$, obtenemos una vez más la ecuación de continuidad. La otra ecuación que obtengo nos dice cómo varía θ ; es

$$\hbar \frac{\partial \theta}{\partial t} = -\frac{m}{2} v^2 - q\phi + \frac{\hbar^2}{2m} \left\{ \frac{1}{\sqrt{\rho}} \nabla^2 (\sqrt{\rho}) \right\}. \quad (21.33)$$

Los que están muy familiarizados con hidrodinámica (de los cuales creo que hay pocos aquí) reconocerán que ésta es la ecuación de movimiento para un fluido cargado eléctricamente si identificamos $\hbar \theta$ con el “potencial de velocidad” —excepto que el último término, que debería ser la energía de compresión del fluido, tiene una dependencia de la densidad ρ —. De cualquier modo, la ecuación nos dice que la derivada de la cantidad $\hbar \theta$ respecto al tiempo está dada por un término de energía cinética, $-\frac{1}{2} m v^2$, más un término de energía potencial, $-q\phi$, con un término adicional que contiene el factor $\hbar^2$, que podríamos llamar “energía cuántica”. Hemos visto que dentro de un superconductor ρ se conserva muy uniforme mediante las fuerzas electrostáticas, por lo que este término se puede despreciar casi con certeza en toda aplicación práctica, siempre que tengamos solamente una región superconductora. Si tenemos un límite entre dos superconductores (u otra circunstancia en que el valor de ρ pueda cambiar rápidamente), este término puede llegar a ser importante.

Para los que no están familiarizados con las ecuaciones de la hidrodinámica, puedo volver a escribir la ecuación (21.33) de una manera que haga más patente lo físico, usando la ecuación (21.31) para expresar θ en función de $\mathbf{v}$. Hallando el gradiente de la ecuación (21.33) y expresando $\nabla \theta$ en función de $\mathbf{A}$ y de $\mathbf{v}$, mediante la ecuación (21.31), obtengo

$$\frac{\partial \mathbf{v}}{\partial t} = \frac{q}{m} \left(-\nabla \phi - \frac{\partial \mathbf{A}}{\partial t} \right) - \mathbf{v} \times (\nabla \times \mathbf{v}) - (\mathbf{v} \cdot \nabla) \mathbf{v} + \nabla \frac{\hbar^2}{2m^2} \left(\frac{1}{\sqrt{\rho}} \nabla^2 \sqrt{\rho} \right). \quad (21.34)$$

¿Qué significa esta ecuación? En primer lugar, recuerden que

$$-\nabla \phi - \frac{\partial \mathbf{A}}{\partial t} = \mathbf{E}. \quad (21.35)$$

Luego noten que, si calculo el rotacional de la ecuación (21.31), obtengo

$$\nabla \times \mathbf{v} = -\frac{q}{m} \nabla \times \mathbf{A}, \quad (21.36)$$

ya que el rotacional de un gradiente da siempre cero. Pero $\nabla \times \mathbf{A}$ es el campo magnético $\mathbf{B}$, por lo que podemos escribir los dos primeros términos en la forma

$$\frac{q}{m} (\mathbf{E} + \mathbf{v} \times \mathbf{B}).$$

Finalmente, deben entender que $\partial \mathbf{v} / \partial t$ representa la derivada respecto al tiempo de la velocidad del fluido en un punto. Si se fijan en una sola partícula, su aceleración es la derivada *total* de $\mathbf{v}$ (o, como se le llama en dinámica de fluidos, la “aceleración sustancial”), que se relaciona con $\partial \mathbf{v} / \partial t$ mediante¹³

$$\left. \frac{d\mathbf{v}}{dt} \right|_{\text{sustancial}} = \frac{\partial \mathbf{v}}{\partial t} + (\mathbf{v} \cdot \nabla) \mathbf{v}. \quad (21.37)$$

¹³ Véase volumen II, sección 40-2.

Este término adicional también aparece como un tercer término en el segundo miembro de la ecuación (21.34). Pasándolo al primer miembro, puedo escribir la ecuación (21.34) como sigue:

$$m \left. \frac{d\mathbf{v}}{dt} \right|_{\text{sustancial}} = q(\mathbf{E} + \mathbf{v} \times \mathbf{B}) + \nabla \frac{\hbar^2}{2m^2} \left(\frac{1}{\sqrt{\rho}} \nabla^2 \sqrt{\rho} \right). \quad (21.38)$$

Según la ecuación (21.36) también tenemos que

$$\nabla \times \mathbf{v} = -\frac{q}{m} \mathbf{B}. \quad (21.39)$$

Estas dos ecuaciones son las ecuaciones de movimiento de un fluido electrónico superconductor. La primera ecuación es simplemente la ley de Newton para un fluido cargado situado en un campo magnético. Dice que la aceleración de cada una de las partículas del fluido cuya carga es q proviene de la fuerza de Lorentz ordinaria $q(\mathbf{E} + \mathbf{v} \times \mathbf{B})$ más una fuerza adicional, que es el gradiente de algún potencial cuántico misterioso —una fuerza que no es muy grande excepto en la juntura de dos superconductores—. La segunda ecuación dice que el fluido es “ideal” —el rotacional de $\mathbf{v}$ da divergencia nula (la divergencia de $\mathbf{B}$ es siempre nula). Esto significa que la velocidad se puede obtener de un potencial de velocidades. Ordinariamente se escribe que $\nabla \times \mathbf{v} = 0$ para un fluido ideal, pero para un *fluido ideal cargado en un campo magnético*, esto se modifica tal como en la ecuación (21.39).

Así, la ecuación de Schrödinger para los pares electrónicos de un superconductor nos da las ecuaciones de movimiento de un fluido ideal cargado eléctricamente. La superconductividad se parece al problema de un líquido cargado en hidrodinámica. Si quieren resolver cualquier problema de superconductores, toman estas ecuaciones para el fluido [o el par equivalente, ecuaciones (21.32) y (21.33)] y las combinan con las ecuaciones de Maxwell para obtener los campos (las cargas y corrientes que usen para obtener los campos deben incluir, naturalmente, las cargas y corrientes del superconductor y de las fuentes externas).

Incidentalmente, creo que la ecuación (21.38) no es muy correcta y que debería tener un término adicional que incluyera la densidad. Este nuevo término no depende de la mecánica cuántica, sino que proviene de la energía ordinaria asociada con variaciones de densidad. Del mismo modo que en un fluido ordinario, debe haber una densidad de energía potencial proporcional al cuadrado de la desviación de ρ respecto a ρ_0 , la densidad no perturbada (que aquí también es igual a la densidad de carga de la red cristalina). Como habrá fuerzas proporcionales al gradiente de esta energía, debe haber otro término en la ecuación (21.38) de la forma: $(\text{const}) \nabla(\rho - \rho_0)^2$. Este término no apareció en el análisis porque proviene de interacciones entre partículas, y desprecié dichas interacciones al usar la aproximación de partículas independientes. Sin embargo, es la fuerza a que me referí cuando hice el enunciado cualitativo de que las fuerzas electrostáticas tienden a conservar ρ casi constante en el interior de un superconductor.

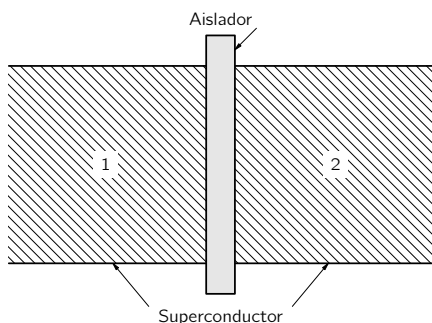


Figura 21-6. Dos superconductores separados por un aislante fino.

21-9 Juntura de Josephson

Ahora quisiera discutir una situación muy interesante que Josephson¹⁴ observó al analizar lo que podía suceder en una juntura entre dos superconductores. Supongan que tenemos dos superconductores unidos mediante una delgada capa de material aislador como se muestra en la figura 21-6. Dicho arreglo se llama ahora “juntura de Josephson”. Si la capa aisladora es gruesa, los electrones no pueden atravesarla; pero si es lo suficientemente delgada, puede haber una amplitud cuántica apreciable de que los electrones pasen. Éste es otro ejemplo de penetración cuántica de una barrera. Josephson analizó esta situación y descubrió que podía ocurrir una cantidad de fenómenos raros.

Para analizar esa juntura, llamaré ψ_1 a la amplitud de encontrar un electrón de un lado y ψ_2 a la de encontrarlo del otro lado. En el estado superconductor,

¹⁴ B. D. Josephson, *Physics Letters* **1**, 251 (1962).

la función de onda ψ_1 es la función de onda corriente de todos los electrones de un lado y ψ_2 es la función correspondiente del otro lado. Podría resolver este problema para diferentes clases de superconductores, pero tomemos una situación muy simple en la cual el material es el mismo de ambos lados, de modo que la juntura es simétrica y sencilla. También supondremos, por un momento, que no existe campo magnético. Siendo así, las dos amplitudes se deben relacionar en la forma siguiente:

$$\begin{aligned} i\hbar \frac{\partial \psi_1}{\partial t} &= U_1 \psi_1 + K \psi_2, \\ i\hbar \frac{\partial \psi_2}{\partial t} &= U_2 \psi_2 + K \psi_1. \end{aligned}$$

La constante K es una característica de la juntura. Si K fuese nula, estas dos ecuaciones describirían simplemente el estado de energía más bajo —con energía U — de cada superconductor. Pero hay un acoplamiento entre los dos lados por medio de la amplitud K de que pueda haber filtración de un lado al otro (es simplemente la amplitud de “ir y venir” de un sistema de dos estados). Si los dos lados son idénticos, U_1 igualaría U_2 y podríamos restar y eliminarlos. Pero supongan que conectamos ahora las dos regiones superconductoras a los dos terminales de una batería para que haya una diferencia de potencial V en la juntura. Entonces, $U_1 - U_2 = qV$. Puedo definir el cero de energía, por conveniencia, como la mitad de esa cantidad y las dos ecuaciones se convierten en

$$\begin{aligned} i\hbar \frac{\partial \psi_1}{\partial t} &= \frac{qV}{2} \psi_1 + K \psi_2, \\ i\hbar \frac{\partial \psi_2}{\partial t} &= -\frac{qV}{2} \psi_2 + K \psi_1. \end{aligned} \tag{21.40}$$

Éstas son las ecuaciones patrón para dos estados cuánticos acoplados. Esta vez vamos a analizar ecuaciones de otro modo. Hagamos las sustituciones

$$\begin{aligned} \psi_1 &= \sqrt{\rho_1} e^{i\theta_1}, \\ \psi_2 &= \sqrt{\rho_2} e^{i\theta_2}, \end{aligned} \tag{21.41}$$

donde θ_1 y θ_2 son las fases a los dos lados de la juntura y ρ_1 y ρ_2 son las densidades de electrones en esos dos puntos. Recuerden que en la práctica ρ_1 y ρ_2 son casi exactamente iguales entre sí y a ρ_0 , la densidad normal de electrones en el material superconductor. Ahora, si sustituyen estas ecuaciones para ψ_1 y ψ_2 en la ecuación (21.40), obtienen cuatro ecuaciones al igualar en cada caso las partes real e imaginaria. Haciendo $(\theta_2 - \theta_1) = \delta$, para abreviar, el resultado es

$$\begin{aligned} \dot{\rho}_1 &= +\frac{2}{\hbar} K \sqrt{\rho_2 \rho_1} \sin \delta, \\ \dot{\rho}_2 &= -\frac{2}{\hbar} K \sqrt{\rho_2 \rho_1} \sin \delta, \end{aligned} \tag{21.42}$$

$$\begin{aligned} \dot{\theta}_1 &= -\frac{K}{\hbar} \sqrt{\frac{\rho_2}{\rho_1}} \cos \delta - \frac{qV}{2\hbar}, \\ \dot{\theta}_2 &= -\frac{K}{\hbar} \sqrt{\frac{\rho_1}{\rho_2}} \cos \delta + \frac{qV}{2\hbar}. \end{aligned} \tag{21.43}$$

Las dos primeras ecuaciones nos dicen que $\dot{\rho}_1 = -\dot{\rho}_2$. “Pero”, dirán, “deben ser las dos cero si ρ_1 y ρ_2 son constantes e iguales a ρ_0 ”. No con tanta prisa. Estas ecuaciones no lo son todo. Dicen lo que $\dot{\rho}_1$ y $\dot{\rho}_2$ serían *si no hubiese fuerzas eléctricas adicionales* debidas a un desequilibrio entre el fluido electrónico y el fondo de iones positivos. Dicen cómo *empezarían* a variar las densidades y, por

consiguiente, describen la clase de corriente que comenzaría a circular. Esta corriente desde el lado 1 hacia el 2 sería simplemente $\dot{\rho}_1$ (o $-\dot{\rho}_2$), es decir,

$$J = \frac{2K}{\hbar} \sqrt{\rho_1 \rho_2} \sin \delta. \quad (21.44)$$

Esa corriente pronto cargaría el lado 2, pero hemos *olvidado* que los dos lados están conectados por alambres a la batería. La corriente que circula no cargará la región 2 (o sea, no descargará la región 1) ya que se establecerán corrientes para mantener constante el potencial. Estas corrientes de baterías no han sido incluidas en nuestras ecuaciones. Si las incluimos, ρ_1 y ρ_2 no varían en realidad, pero aun así la corriente a través de la juntura está dada por la ecuación (21.44).

Como ρ_1 y ρ_2 permanecen constantes e iguales a ρ_0 , pongamos $2K\rho_0/\hbar = J_0$, y escribamos

$$J = J_0 \sin \delta. \quad (21.45)$$

J_0 , al igual que K , es sólo un número característico de la juntura particular.

El otro par de ecuaciones de (21.43) nos hablan de θ_1 y θ_2 . Nos interesa la diferencia $\delta = \theta_2 - \theta_1$ para usar la ecuación (21.45); lo que obtenemos es

$$\dot{\delta} = \dot{\theta}_2 - \dot{\theta}_1 = \frac{qV}{\hbar}. \quad (21.46)$$

Esto significa que podemos escribir

$$\delta(t) = \delta_0 + \frac{q}{\hbar} \int V(t) dt, \quad (21.47)$$

donde δ_0 es el valor de δ en $t = 0$. Recuerden también que q es la carga de un par, es decir $q = 2q_e$. Tenemos un resultado importante en las ecuaciones (21.45) y (21.47), la teoría general de la juntura de Josephson.

Pero ¿cuales son las consecuencias? En primer lugar, apliquen un voltaje V_0 de CC. Si lo hacen, el argumento del seno, es $(\delta_0 + (q/\hbar)V_0 t)$. Como $\hbar$ es un número pequeño (comparado con voltajes y tiempos ordinarios), el seno oscila muy rápidamente y la corriente resultante es nula (en la práctica, como la temperatura es distinta de cero, obtendrían una corriente pequeña que se debe a la conducción por electrones “normales”). Por otro lado, si tiene un voltaje *nulo* en la juntura, ¡pueden obtener una corriente! Sin voltaje la corriente puede tener cualquier calor comprendido entre $+J_0$ y $-J_0$ (dependiendo esto del valor de δ_0). Pero traten de establecer un voltaje y la corriente va a cero. Este comportamiento extraño se ha observado en experimentos recientes.¹⁵

Hay otro modo de obtener una corriente —aplicando un voltaje de muy alta frecuencia además del voltaje CC—. Hagamos

$$V = V_0 + v \cos \omega t,$$

donde $v \ll V$. Así, $\delta(t)$ es

$$\delta_0 + \frac{q}{\hbar} V_0 t + \frac{q}{\hbar} \frac{v}{\omega} \sin \omega t.$$

Ahora bien, para Δx pequeños,

$$\sin(x + \Delta x) \approx \sin x + \Delta x \cos x.$$

Usando esta aproximación de $\sin \delta$, obtengo

$$J = J_0 \left[\sin \left(\delta_0 + \frac{q}{\hbar} V_0 t \right) + \frac{q}{\hbar} \frac{v}{\omega} \sin \omega t \cos \left(\delta_0 + \frac{q}{\hbar} V_0 t \right) \right].$$

El promedio del primer término es cero, pero el del segundo término no lo es si

$$\omega = \frac{q}{\hbar} V_0.$$

¹⁵ P. W. Anderson y J. M. Rowell, *Phys. Rev. Letters* **10**, 230 (1963).

Debe haber una corriente si el voltaje CA tiene justamente esta frecuencia. Shapiro¹⁶ sostiene haber observado este efecto de resonancia.

Si revisan los trabajos sobre este tema, encontrarán que escriben con frecuencia la fórmula para la corriente en la forma

$$J = J_0 \sin\left(\delta_0 + \frac{2q_e}{\hbar} \int \mathbf{A} \cdot d\mathbf{s}\right), \quad (21.48)$$

donde la integral se ha de efectuar a través de la juntura. La razón de ello es que cuando hay un potencial vectorial en la juntura, la fase de amplitud del ir y venir se modifica como explicamos anteriormente. Si hallan esta fase adicional resulta que es la que dimos antes.

Finalmente, me gustaría describir un experimento muy espectacular e interesante que se ha llevado a cabo recientemente sobre la interferencia de las corrientes provenientes de las dos junturas. En mecánica cuántica estamos acostumbrados a la interferencia entre amplitudes provenientes de dos rendijas diferentes. Ahora vamos a tener interferencia entre dos junturas, causada por el defasaje a la llegada de corrientes que siguen dos caminos diferentes. En la figura 21-7 muestro dos junturas “a” y “b” diferentes conectadas en paralelo. Los terminales P y Q se conectan a nuestros instrumentos eléctricos que miden cualquier flujo de corriente. La corriente externa, J_{total} , será la suma de las corrientes a través de las dos junturas. Sean J_a y J_b estas dos corrientes y δ_a y δ_b sus fases respectivas. Ahora bien, el defasaje de las funciones de onda entre P y Q debe ser el mismo, ya vayan por un camino u otro. En el camino a través de la juntura “a” la diferencia de fase entre P y Q es δ_a más la integral de línea del potencial vectorial a lo largo del camino superior:

$$\Delta\text{Fase}_{P \rightarrow Q} = \delta_a + \frac{2q_e}{\hbar} \int_{\text{superior}} \mathbf{A} \cdot d\mathbf{s}. \quad (21.49)$$

¿Por qué? Porque la fase θ se relaciona con $\mathbf{A}$ mediante la ecuación (21.26). Si integran dicha ecuación mediante algún medio, el primer miembro da la variación de fase, que es exactamente proporcional a la integral de línea $\mathbf{A}$, como acabamos de escribir. La variación de fase según el camino inferior se puede escribir de un modo análogo

$$\Delta\text{Fase}_{P \rightarrow Q} = \delta_b + \frac{2q_e}{\hbar} \int_{\text{inferior}} \mathbf{A} \cdot d\mathbf{s}. \quad (21.50)$$

Las dos deben ser iguales; si las resto obtengo que la diferencia de las deltas debe ser la integral de línea de $\mathbf{A}$ a lo largo del circuito:

$$\delta_b - \delta_a = \frac{2q_e}{\hbar} \oint_{\Gamma} \mathbf{A} \cdot d\mathbf{s}.$$

Aquí la integral es según el camino cerrado Γ de la figura 21-7 que pasa por las dos junturas. La integral sobre $\mathbf{A}$ es el flujo magnético Φ a través del lazo. Así, pues, la diferencia entre las δ será $2q_e/\hbar$ por el flujo magnético Φ que pasa entre las dos ramas del circuito:

$$\delta_b - \delta_a = \frac{2q_e}{\hbar} \Phi. \quad (21.51)$$

Puedo controlar esta diferencia de fase cambiando el campo magnético en el circuito, por lo que puedo ajustar el defasaje y ver si la corriente total circula por las dos junturas mostrando o no cualquier interferencia debida a las dos partes. La corriente total será la suma de J_a y J_b . Escribiré, por conveniencia

$$\delta_a = \delta_0 + \frac{q_e}{\hbar} \Phi, \quad \delta_b = \delta_0 - \frac{q_e}{\hbar} \Phi.$$

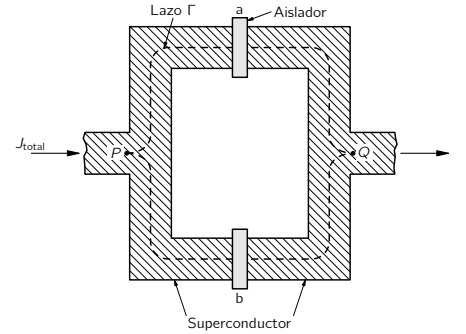


Figura 21-7. Dos junturas de Josephson en paralelo.

¹⁶ S. Shapiro, *Phys. Rev. Letters* **11**, 80 (1963).

Entonces,

$$J_{\text{total}} = J_0 \left\{ \sin \left(\delta_0 + \frac{q_e}{\hbar} \Phi \right) + \sin \left(\delta_0 - \frac{q_e}{\hbar} \Phi \right) \right\} \\ = 2J_0 \sin \delta_0 \cos \frac{q_e \Phi}{\hbar}. \quad (21.52)$$

Pero no sabemos nada de δ_0 y la naturaleza puede ajustar eso como quiera, dependiendo de las circunstancias. En particular, dependerá del voltaje externo que apliquemos a la juntura. Sin embargo, hagamos lo que hagamos, $\sin \delta_0$ nunca será mayor que 1. Por lo que la corriente *máxima* para cualquier Φ está dada por

$$J_{\text{máx}} = 2J_0 \left| \cos \frac{q_e \Phi}{\hbar} \right|.$$

Esta corriente máxima variará con Φ y tendrá máximos siempre que

$$\Phi = n \frac{\pi \hbar}{q_e},$$

con n un número entero. Es decir, ¡la corriente alcanza sus valores máximos donde el flujo tiene justamente los valores cuantizados que encontramos en la ecuación (21.30)!

La corriente de Josephson a través de una juntura doble se ha medido recientemente¹⁷ en función del campo magnético en el área entre las junturas. La figura 21-8 muestra los resultados. Hay un fondo general de corriente proveniente de diversos efectos que hemos despreciado, pero las oscilaciones rápidas de la corriente con variaciones en el campo magnético se deben al término interferencial $\cos q_e \Phi / \hbar$ de la ecuación (21.52).

Uno de los puntos de la mecánica cuántica que despiertan interés es la cuestión de si existe potencial vectorial en un lugar donde no existe campo.¹⁸ El experimento que acabo de mencionar se ha realizado también con un solenoide pequeñísimo entre las dos junturas de modo que el único campo magnético $\mathbf{B}$ significativo es el que está dentro del solenoide y una cantidad despreciable se encuentra en los hilos superconductores. Se ha dicho también que la cantidad de corriente depende, oscilatoriamente, del flujo del campo magnético dentro de dicho solenoide, aunque ese campo nunca toca los alambres —otra demostración de la “realidad física” del potencial vectorial—. ¹⁹

No sé qué surgirá luego, pero miren lo que se puede hacer por ahora. En primer lugar, observen que la interferencia entre dos junturas se puede usar para construir un magnetómetro sensible. Si se construye un par de junturas que encierran un área de 1 mm^2 , digamos, los máximos de la curva de la figura 21-8 estarían separados 2×10^{-6} gauss. Ciertamente es posible saber cuándo se

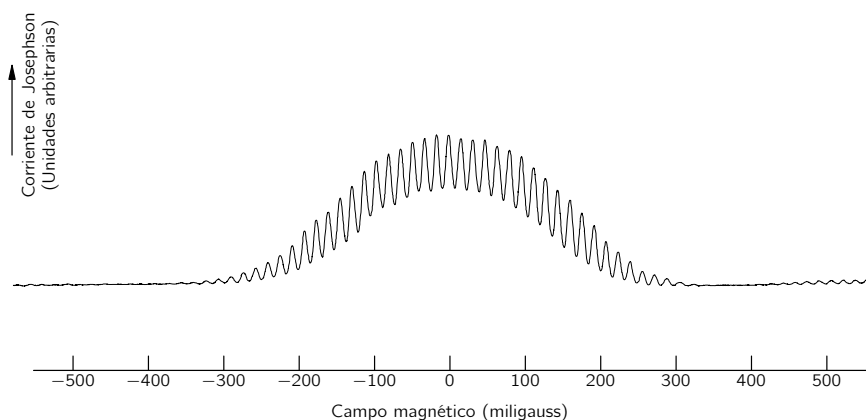


Figura 21-8. Una grabación de la corriente a través de un par de junturas de Josephson como una función del campo magnético en la región entre las dos junturas (véase figura 21-7). [Esta grabación se debe a R. C. Jaklevic, J. Lambe, A. H. Silver y J. E. Mercereau de Scientific Laboratory, Ford Motor Company.]

¹⁷ Jaklevic, Lambe, Silver y Mercereau, *Phys. Rev. Letters* **12**, 159 (1964).

¹⁸ Jaklevic, Lambe, Silver y Mercereau, *Phys. Rev. Letters* **12**, 274 (1964).

¹⁹ Véase volumen II, capítulo 15, sección 15-5.

encuentran a $1/10$ entre dos picos; por ello sería posible usar una juntura como ésa para medir campos magnéticos del orden de 2×10^{-7} gauss —o para medir campos mayores con esa precisión—. Se podría avanzar aún más. Supongan, por ejemplo, que colocamos un conjunto de 10 o 20 junturas muy próximas y con igual espaciamiento. Podemos tener así la interferencia entre 10 o 20 rendijas y a medida que variemos el campo magnético obtendremos máximos y mínimos muy agudos. En lugar de una interferencia por dos rendijas, podemos tener un interferómetro de 20 o, quizás aún, de 100 rendijas para medir el campo magnético. Quizá podemos predecir que la medida de campos magnéticos llegará a ser con el tiempo —usando los efectos de interferencia cuántica— casi tan precisos como la medida de la longitud de onda de la luz.

Éstas son algunas de las cosas que están sucediendo actualmente —el transistor, el láser— y ahora estas junturas de las cuales no se conoce aún todas las aplicaciones prácticas. La mecánica cuántica, descubierta en 1926, ha tenido casi 40 años de desarrollo y de repente ha comenzado a ser explotada de muchas maneras prácticas y reales. Realmente estamos controlando la naturaleza a un nivel muy hermoso y sutil.

Lamento decirles a todos que, para participar en esta aventura, es un imperativo absoluto que ustedes aprendan mecánica cuántica lo más pronto posible. Nuestra esperanza era que en este curso encontraríamos el modo de hacerles comprender lo más tempranamente posible los misterios de esta parte de la física.

Epílogo de Feynman

Bueno, he estado hablando con ustedes dos años y ahora voy a dejarlos. En cierto sentido quisiera disculparme y en otro no. Espero —en realidad sé— que dos o tres docenas de ustedes han podido seguir todo con gran entusiasmo y que han pasado muy buenos ratos. Pero también sé que “los poderes de la instrucción son muy poco eficaces, excepto en aquellas felices circunstancias en que son prácticamente superfluos”. Así, pues, a esas dos o tres docenas que comprendieron todo, podría decirles que yo no hice nada más que mostrarles las cosas. En cuanto a los demás, si los he hecho odiar la materia, lo siento. Nunca había enseñado física elemental hasta ahora y les pido disculpas. Lo único que deseo es que no les haya ocasionado grandes inconvenientes y que ustedes no se aparten de este quehacer tan interesante. Espero que alguien pueda enseñarles lo mismo de un modo que no les dé indigestión y que algún día descubran que, después de todo, esto no es tan horrible como parece. Por último, permítanme agregar que la intención principal de mis clases no ha sido prepararlos para un examen —tampoco prepararlos para trabajar en la industria o en las fuerzas armadas—. El propósito mayor ha sido hacerles apreciar lo maravilloso que es el mundo y cómo lo encara el físico, porque creo sinceramente que esto constituye una gran parte de la verdadera cultura de los tiempos modernos (hay profesores de otras materias que probablemente lo objetarán, pero en mi opinión ellos están completamente equivocados). Tal vez ustedes no sólo hayan llegado a la valoración de este aspecto de la cultura; quizá quieran participar también en la aventura más grandiosa que jamás haya emprendido la mente humana.

Apéndice

Gran parte del trabajo de este volumen presupone un conocimiento del tema magnetismo atómico que se trató en los capítulos 34 y 35 del volumen II. Para comodidad de los lectores que no tengan el volumen II a mano, se han reproducido aquí esos dos capítulos.

Su contenido es el siguiente:

Capítulo 34 Magnetismo de la materia

- 34-1 Diamagnetismo y paramagnetismo
- 34-2 Momentos magnéticos y momentum angular
- 34-3 Precesión de los imanes atómicos
- 34-4 Diamagnetismo
- 34-5 Teorema de Larmor
- 34-6 La física clásica no da ni diamagnetismo ni paramagnetismo
- 34-7 Momentum angular en la mecánica cuántica
- 34-8 Energía magnética de los átomos

Capítulo 35 Paramagnetismo y resonancias magnéticas

- 35-1 Estados magnéticos cuantizados
- 35-2 El experimento de Stern-Gerlach
- 35-3 El método de haces moleculares de Rabi
- 35-4 Paramagnetismo macroscópico de materiales
- 35-5 Enfriamiento por desmagnetización adiabática
- 35-6 Resonancia magnética nuclear

Magnetismo de la materia

34-1 Diamagnetismo y paramagnetismo

En este capítulo vamos a hablar de las propiedades magnéticas de los materiales. El material que tiene las propiedades magnéticas más sorprendentes es, por supuesto, el hierro. Los elementos níquel, cobalto y —a temperaturas suficientemente bajas (por debajo de 16°C)— gadolinio, así como cierta cantidad de aleaciones especiales, tienen propiedades magnéticas similares. Esa clase de magnetismo, llamada *ferromagnetismo*, que es tan sorprendente y complicado, lo estudiaremos en un capítulo especial. Sin embargo, todas las sustancias ordinarias muestran algunos efectos magnéticos, aunque muy pequeños —mil a un millón de veces menores que los efectos de los materiales ferromagnéticos—. Aquí vamos a describir el magnetismo común, es decir, el magnetismo de las sustancias que no son ferromagnéticas.

Este pequeño magnetismo es de dos clases. Algunos materiales son *atraídos* hacia campos magnéticos; otros son *repelidos*. Contrariamente al efecto eléctrico en la materia, que siempre hace que los dieléctricos sean atraídos, el efecto magnético tiene dos signos. Es posible mostrar estos dos signos con ayuda de un electroimán intenso que tiene una pieza polar muy afilada y una pieza polar chata, como se ha dibujado en la figura 34-1. El campo magnético es mucho más intenso cerca del polo en punta que cerca del polo chato. Si se ata un pequeño trozo de material a un hilo largo y se suspende entre los polos, habrá en general una pequeña fuerza sobre él. Se puede ver esta pequeña fuerza por el ligero desplazamiento del material colgado cuando se enciende el imán. Los pocos materiales ferromagnéticos son atraídos muy fuertemente hacia el polo en punta; todos los otros materiales experimentan una fuerza muy débil. Algunos son atraídos débilmente hacia el polo en punta; otros son repelidos débilmente.

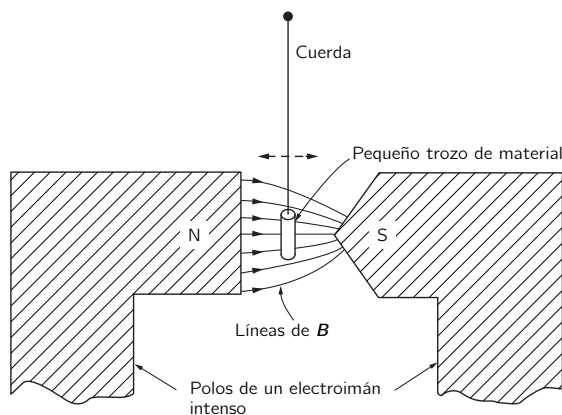


Figura 34-1. El polo en punta repele débilmente un pequeño cilindro de bismuto y atrae un trozo de aluminio.

La manera más fácil de ver el efecto es con un cilindro pequeño de bismuto, que es *repelido* de la región de campo intenso. Las sustancias repelidas de este modo se llaman *diamagnetismo*. El bismuto es uno de los materiales diamagnéticos más intensos, pero aun así el efecto es bastante débil. El diamagnetismo siempre es muy débil. Si se suspende un pedazo pequeño de aluminio entre los polos, también hay una fuerza débil, pero *hacia* el polo en punta. Las sustancias como el aluminio se llaman *paramagnetismo*. (En un experimento como éste aparecen

- 34-1 Diamagnetismo y paramagnetismo
- 34-2 Momentos magnéticos y momentum angular
- 34-3 Precesión de los imanes atómicos
- 34-4 Diamagnetismo
- 34-5 Teorema de Larmor
- 34-6 La física clásica no da ni diamagnetismo ni paramagnetismo
- 34-7 Momentum angular en la mecánica cuántica
- 34-8 Energía magnética de los átomos

Revisar

Volumen II, sección 15-1,
“Las fuerzas sobre un lazo de corriente; energía de un dipolo”

fuerzas debidas a corrientes de Foucault cuando se enciende y se apaga el imán, las cuales pueden dar lugar a impulsos fuertes. Deben tener cuidado de determinar el desplazamiento neto después que el objeto colgado se queda quieto.)

Ahora queremos describir brevemente los mecanismos de estos dos efectos. Primero, en muchas sustancias los átomos no tienen momento magnético permanente, o más bien, todos los imanes que hay dentro de cada átomo se compensan de modo que el momento *neto* del átomo es cero. Los espines y los movimientos orbitales electrónicos se compensan exactamente, de modo que cualquier átomo particular no tiene momento magnético medio. En estas circunstancias, cuando aplican un campo magnético se generan por inducción pequeñas corrientes adicionales dentro del átomo. Según la ley de Lenz, el sentido de estas corrientes es tal que se oponen al aumento del campo. Así, los momentos magnéticos inducidos de los átomos están en dirección *opuesta* al campo magnético. Éste es el mecanismo del diamagnetismo.

Luego hay algunas sustancias cuyos átomos sí tienen un momento magnético permanente —átomos en los cuales las órbitas y los espines electrónicos tienen una corriente neta no nula—. Así, además del efecto diamagnético (que siempre está presente), está la posibilidad de alinear los momentos magnéticos atómicos. En este caso, los momentos tratan de alinearse *con* el campo magnético (como los dipolos permanentes de un dieléctrico se alinean con el campo eléctrico), y el magnetismo inducido tiende a aumentar el campo magnético. Éstas son las sustancias paramagnéticas. Generalmente el paramagnetismo es bastante débil porque las fuerzas de alineamiento son relativamente bajas frente a las fuerzas provenientes de los movimientos térmicos que tienden a destruir el orden. Es por eso también que el paramagnetismo es normalmente sensible a la temperatura. (El paramagnetismo que proviene de los espines de los electrones responsables de la conducción en un metal constituye la excepción. No estudiaremos este fenómeno aquí.) En el paramagnetismo ordinario, cuanto menor es la temperatura mayor es el efecto. Hay más alineamientos a temperaturas bajas cuando los efectos de desordenamiento debido a las colisiones son menores. Por el contrario, el diamagnetismo es más o menos independiente de la temperatura. En cualquier sustancia con momentos magnéticos intrínsecos hay tanto un efecto diamagnético como uno paramagnético, pero domina comúnmente este último.

En el capítulo 11 describimos un material *ferroeléctrico* en el que todos los dipolos eléctricos se alineaban debido a sus propios campos eléctricos mutuos. También es posible imaginar el análogo magnético de la ferroelectricidad, en el cual todos los momentos se alinearían y se acoplarían. Si hacen cálculos sobre cómo ocurriría esto, encontrarán que como las fuerzas magnéticas son mucho más pequeñas que las eléctricas, los movimientos térmicos destruirán este alineamiento aun a temperaturas tan bajas como unas décimas de grado Kelvin. Así, sería imposible a la temperatura ambiente tener un alineamiento permanente de los imanes.

Por otra parte, esto es exactamente lo que ocurre en el hierro —ahí sí hay alineamiento—. Hay una fuerza efectiva entre los momentos magnéticos de diferentes átomos de hierro, la cual es mucho, mucho mayor que la interacción *magnética directa*. Es un efecto indirecto que sólo se puede explicar con la mecánica cuántica. Es unas 10 000 veces más fuerte que la interacción magnética directa y es lo que alinea los momentos en los materiales ferromagnéticos. Estudiamos esta interacción especial en un capítulo posterior.

Ahora que hemos tratado de darles una explicación cualitativa del diamagnetismo y del paramagnetismo, debemos corregirnos y decir honestamente que *no es posible* comprender los efectos magnéticos de los materiales desde el punto de vista de la física clásica. Estos efectos magnéticos son *fenómenos completamente cuánticos*. Sin embargo, es posible hacer ciertos razonamientos clásicos engañosos y hacerse una idea de lo que está pasando. Digámoslo así. Pueden hacer razonamientos clásicos y obtener estimaciones sobre el comportamiento del material, pero estos razonamientos no son “legales” en ningún sentido porque es absolutamente esencial hacer intervenir a la mecánica cuántica en cada uno de estos fenómenos magnéticos. No obstante, hay situaciones, como en un plasma

o región del espacio con muchos electrones libres, en las que los electrones sí obedecen las leyes de la mecánica clásica. Y en estas circunstancias tienen valor algunos teoremas del magnetismo clásico. Adicionalmente, los razonamientos clásicos tienen cierto valor por razones históricas. Las pocas primeras veces que los físicos pudieron hacerse una idea del significado y comportamiento de los materiales magnéticos, utilizaron razonamientos clásicos. Finalmente, y ya hemos dado ejemplos de esto, la mecánica clásica nos puede dar estimaciones útiles de lo que está pasando —aunque la manera realmente honesta de estudiar este tema sería aprendiendo mecánica cuántica primero para luego entender el magnetismo en términos de la mecánica cuántica—.

Por otra parte, no queremos esperar hasta que aprendamos mecánica cuántica de punta a punta para comprender algo simple como el diamagnetismo. Tendremos que descansar en la mecánica clásica en una especie de demostración a medias de lo que ocurre, teniendo en cuenta, sin embargo, que los razonamientos no son realmente correctos. Por lo tanto, hacemos una serie de teoremas sobre el magnetismo clásico, que los confundirán porque demostrarán cosas diferentes. A excepción del último, todos serán incorrectos. Aun más, todos serán incorrectos como descripción del mundo físico porque se deja fuera a la mecánica cuántica.

34-2 Momentos magnéticos y momentum angular

El primer teorema que queremos demostrar a partir de la mecánica clásica es el siguiente: si un electrón se mueve en una órbita circular (por ejemplo, dando vueltas alrededor de un núcleo bajo la influencia de una fuerza central), hay un cociente definido entre el momento magnético y el momentum angular. Llamemos $\mathbf{J}$ al momentum angular y $\boldsymbol{\mu}$ al momento magnético del electrón en la órbita. El módulo del momentum angular es la masa del electrón por la velocidad por el radio (véase figura 34-2). Está dirigido perpendicularmente al plano de la órbita.

$$J = mvr. \quad (34.1)$$

(Ésta es, naturalmente, una fórmula no relativista, pero es una buena aproximación para átomos, porque para los electrones que intervienen, v/c es generalmente del orden de $e^2/\hbar c \approx 1/137$ o sea, aproximadamente 1 %.)

El momento magnético de la misma órbita es la corriente por el área. (Véase sección 14-5.) La corriente es la carga que pasa por unidad de tiempo por cualquier punto de la órbita, o sea, la carga q por la frecuencia de rotación. La frecuencia es la velocidad dividida entre la longitud de la órbita; luego,

$$I = q \frac{v}{2\pi r}.$$

El área es πr^2 , de modo que el momento magnético es

$$\mu = \frac{qvr}{2}. \quad (34.2)$$

También es perpendicular al plano de la órbita. Por lo tanto, $\mathbf{J}$ y $\boldsymbol{\mu}$ están en la misma dirección

$$\boldsymbol{\mu} = \frac{q}{2m} \mathbf{J} \text{ (órbita)}. \quad (34.3)$$

Su cociente no depende ni de la velocidad ni del radio. Para cualquier partícula moviéndose en una órbita circular el momento magnético es igual a $q/2m$ por el momentum angular. En el caso del electrón la carga es negativa —la podemos llamar $-q_e$ —; luego, para un electrón,

$$\boldsymbol{\mu} = -\frac{q_e}{2m} \mathbf{J} \text{ (órbita del electrón)}. \quad (34.4)$$

Esto es lo que se esperaría clásicamente y, lo que es más bien un milagro, también vale en la mecánica cuántica. Son cosas que pasan. Sin embargo, si continúan con la física clásica, encontrarán otras partes donde da respuestas incorrectas, y es todo juego tratar de recordar qué cosas están bien y qué cosas

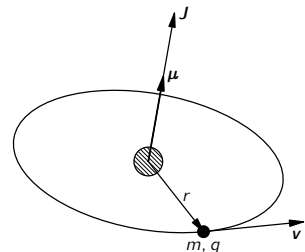


Figura 34-2. Para cualquier órbita circular el momento magnético $\boldsymbol{\mu}$ es $q/2m$ por el momentum angular $\mathbf{J}$.

son incorrectas. También podríamos darles inmediatamente lo que es válido *en general* en la mecánica cuántica. Primero, la ecuación (34.4) es válida para el *movimiento orbital*, pero no es el único magnetismo que existe. El electrón también tiene espín, que es algo así como una rotación alrededor de su propio eje (parecida a la rotación de la Tierra alrededor de su eje) y como resultado de ese espín tiene también un momentum angular y un momento magnético. Pero por razones puramente cuánticas —no hay explicación clásica— el cociente entre $\boldsymbol{\mu}$ y $\mathbf{J}$ para el espín del electrón es el doble del correspondiente al movimiento orbital del electrón con espín:

$$\boldsymbol{\mu} = -\frac{q_e}{m} \mathbf{J} \text{ (espín del electrón)}. \quad (34.5)$$

En cualquier átomo hay, en general, varios electrones y algunas combinaciones de espín y rotación orbital que dan lugar a un momentum angular total y a un momento magnético total. Aunque no hay ninguna razón clásica de por qué debe ser así, *siempre es verdad* en la mecánica cuántica que (para un átomo aislado) la dirección del momento magnético es exactamente opuesta a la dirección del momentum angular. El cociente de los dos no es necesariamente ni $-q_e/m$ ni $-q_e/2m$ sino algo intermedio porque hay una mezcla de las contribuciones de las órbitas y de los espines. Podemos escribir

$$\boldsymbol{\mu} = -g \left(\frac{q_e}{2m} \right) \mathbf{J}, \quad (34.6)$$

donde g es un factor característico del estado del átomo. Sería 1 para un momento puramente orbital, o 2 para un momento puramente de espín, o algo intermedio para un sistema complicado como un átomo. Naturalmente, esta fórmula no nos dice mucho. Indica que el momento magnético es *paralelo* al momentum angular, pero puede tener cualquier módulo. La forma de la ecuación (34.6) es conveniente, porque g —llamado “factor g de Landé”— es una constante adimensional cuya magnitud es de orden de uno. Una de las tareas de la mecánica cuántica es predecir el factor g para cualquier estado atómico particular.

Puede que también estén interesados en lo que pasa en los núcleos. En los núcleos hay protones y neutrones que se pueden mover en una especie de órbita y al mismo tiempo tienen, como el electrón, una rotación intrínseca o espín. Aquí también el momento magnético es paralelo al momentum angular. Sólo que ahora el orden de magnitud del cociente de los dos es lo que sería de esperar para un *protón* dando vueltas en círculo, con la m de la ecuación (34.3) igual a la *masa del protón*. Por lo tanto, se acostumbra a escribir para los núcleos

$$\boldsymbol{\mu} = g \left(\frac{q_e}{2m_p} \right) \mathbf{J}, \quad (34.7)$$

donde m_p es la masa del protón y g —llamado factor g *nuclear*— es un número cercano a uno, que hay que determinar para cada núcleo.

Otra diferencia importante en un núcleo es que el momento magnético de *espín* del protón *no* tiene un factor g de 2 como el electrón. Para un protón, $g = 2(2.79)$. Y lo que es sorprendente, el *neutrón* también tiene un momento magnético de espín y el cociente entre su momento magnético y su momentum angular es $2(-1.91)$. En otras palabras, el neutrón no es exactamente “neutro” en lo que respecta al magnetismo. Es como un pequeño imán y tiene la clase de momento magnético que tendría una carga *negativa* girando.

34-3 Precesión de los imanes atómicos

Una de las consecuencias de tener el momento magnético proporcional al momentum angular es que un imán atómico colocado en campo magnético *precesa*. Primero razonaremos clásicamente. Supongan que tenemos un momento magnético $\boldsymbol{\mu}$ suspendido libremente en un campo magnético uniforme. Experimentará un torque $\boldsymbol{\tau}$, igual a $\boldsymbol{\mu} \times \mathbf{B}$ que trata de alinearlo en la dirección del campo.

Pero el imán atómico es un giroscopio —tiene un momentum angular $\mathbf{J}$ —. Por lo tanto, debido al campo magnético el torque no hará que el imán termine alineado. En lugar de eso, el imán *precesará*, como vimos al analizar un giroscopio en el capítulo 20 del volumen I. El momentum angular —y con él el momento magnético— precesa alrededor de un eje paralelo al campo magnético. Podemos hallar la velocidad de precesión con el mismo método que usamos en el capítulo 20 del primer volumen.

Supongan que en un corto tiempo Δt el momentum angular varía de $\mathbf{J}$ a $\mathbf{J}'$, como se ha dibujado en la figura 34-3, formando siempre el mismo ángulo θ con la dirección del campo magnético $\mathbf{B}$. Llamemos ω_p a la velocidad angular de precesión, por lo que en el tiempo Δt el ángulo de precesión es $\omega_p \Delta t$. Se deduce de la geometría de la figura, que la variación del momentum angular en el tiempo Δt es

$$\Delta \mathbf{J} = (J \sin \theta)(\omega_p \Delta t).$$

Por lo tanto, la derivada del momentum angular respecto al tiempo es

$$\frac{d\mathbf{J}}{dt} = \omega_p J \sin \theta, \quad (34.8)$$

que debe ser igual al torque:

$$\tau = \mu B \sin \theta. \quad (34.9)$$

La velocidad angular de precesión es entonces

$$\omega_p = \frac{\mu}{J} B. \quad (34.10)$$

Sustituyendo μ/J dado en la ecuación (34.6), vemos que para un sistema atómico se tiene

$$\omega_p = g \frac{q_e B}{2m}; \quad (34.11)$$

la frecuencia de precesión es proporcional a B . Es práctico recordar que para un átomo (o un electrón)

$$f_p = \frac{\omega_p}{2\pi} = (1.4 \text{ megaciclos/gauss})gB, \quad (34.12)$$

y que para un núcleo

$$f_p = \frac{\omega_p}{2\pi} = (0.76 \text{ kilociclos/gauss})gB. \quad (34.13)$$

(Las fórmulas para átomos y para núcleos son diferentes únicamente por las convenciones distintas para g en ambos casos.)

Por lo tanto, según la teoría *clásica* las órbitas —y los espines— electrónicas de un átomo deben precesar en un campo magnético. ¿Es también válido en la mecánica cuántica? Esencialmente es válido, pero el significado de la “precesión” es diferente. En la mecánica cuántica no se puede hablar de *dirección* del momentum angular en el mismo sentido que clásicamente; de todas maneras, hay una analogía muy estrecha —tan estrecha que seguiremos llamándola “precesión”—. La estudiaremos más adelante cuando hablemos del punto de vista cuántico.

34-4 Diamagnetismo

A continuación queremos examinar el *diamagnetismo* desde el punto de vista clásico. Se puede calcular de varias maneras, pero una de las formas más elegantes es la siguiente. Supongan que encendemos lentamente un campo magnético en las cercanías de un átomo. A medida que el campo magnético cambia, se genera un campo *eléctrico* por inducción magnética. Conforme a la ley de Faraday, la integral de línea de $\mathbf{E}$ a lo largo de cualquier camino cerrado es la derivada respecto al tiempo del flujo magnético a través del camino. Supongan que elegimos una trayectoria Γ que es una circunferencia de radio r alineada al centro del

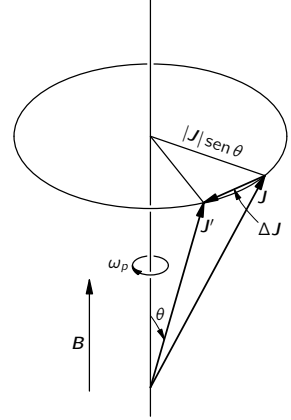


Figura 34-3. Un objeto de momentum angular $\mathbf{J}$ y momento magnético paralelo μ colocado en un campo magnético $\mathbf{B}$ precesa con velocidad angular ω_p .

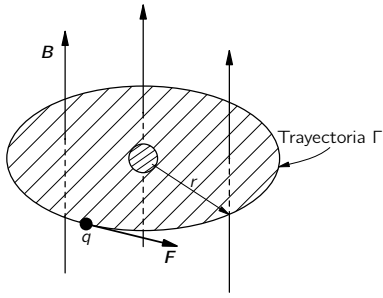


Figura 34-4. Fuerzas eléctricas inducidas que actúan sobre los electrones en un átomo.

átomo, como muestra la figura 34-4. El promedio del campo eléctrico tangencial E a lo largo de este camino está dado por

$$E2\pi r = -\frac{d}{dt}(B\pi r^2),$$

y hay un campo eléctrico circular cuya intensidad es

$$E = -\frac{r}{2} \frac{dB}{dt}.$$

El campo eléctrico inducido que actúa sobre un electrón del átomo produce un torque igual a $-q_e E r$, el cual debe ser igual a la derivada del momentum angular respecto al tiempo, dJ/dt :

$$\frac{dJ}{dt} = \frac{q_e r^2}{2} \frac{dB}{dt}. \quad (34.14)$$

Integrando respecto al tiempo desde el campo cero, encontramos que la variación del momentum angular debida al encendido del campo es

$$\Delta J = \frac{q_e r^2}{2} B. \quad (34.15)$$

Éste es el momentum angular adicional proveniente del impulso rotatorio dado a los electrones al encender el campo.

Este momentum adicional da lugar a un momento magnético adicional que, como se trata de un movimiento *orbital*, es simplemente $-q_e/2m$ por el momentum angular. El momento diamagnético inducido es

$$\Delta\mu = -\frac{q_e}{2m} \Delta J = -\frac{q_e^2 r^2}{4m} B. \quad (34.16)$$

El signo menos (que con la regla de Lenz pueden ver que es correcto) significa que el momento adicional es opuesto al campo magnético.

Nos gustaría escribir la ecuación (34.16) de un modo un poco diferente. El r^2 que aparece es el radio a partir de un eje paralelo a $\mathbf{B}$ que pasa por el átomo, por lo que si $\mathbf{B}$ está según el eje z , es $x^2 + y^2$. Si consideramos los átomos esféricamente simétricos (o promediamos sobre átomos con sus ejes naturales en todas direcciones), el promedio de $x^2 + y^2$ es $2/3$ del promedio del cuadrado de la verdadera distancia radial desde el *punto* central del átomo. En consecuencia, por lo común, es más conveniente escribir la ecuación (34.16) en la forma

$$\Delta\mu = -\frac{q_e^2}{6m} \langle r^2 \rangle_{\text{prom}} B. \quad (34.17)$$

En todo caso, hemos encontrado un momento atómico inducido proporcional al campo magnético B y opuesto a él. Éste es el diamagnetismo de la materia. Es este efecto magnético el responsable de la pequeña fuerza que experimenta un trozo de bismuto en un campo magnético no uniforme (pueden obtener la fuerza calculando la energía de los momentos inducidos en el campo y viendo cómo varía la energía cuando el material se saca o se introduce en la región del campo intenso).

Aún nos queda un problema: ¿cuál es el promedio $\langle r^2 \rangle_{\text{prom}}$ del cuadrado del radio? La mecánica clásica no nos puede dar respuesta alguna. Debemos volver atrás y empezar todo con la mecánica cuántica. En un átomo, no podemos decir realmente dónde está un electrón, sino conocer únicamente la probabilidad de que esté en un sitio. Si interpretamos $\langle r^2 \rangle_{\text{prom}}$ como el promedio del cuadrado de la distancia al centro para la distribución de probabilidad, el momento diamagnético dado por la mecánica cuántica es exactamente lo mismo que la fórmula (34.17). Naturalmente, esta ecuación es el momento de un electrón. El momento total está dado por la suma sobre todos los electrones del átomo. Lo sorprendente es que el razonamiento clásico y el de la mecánica cuántica dan la misma respuesta,

aunque como veremos, el razonamiento clásico que lleva a la ecuación (34.17) no es en realidad válido en la mecánica clásica.

Aparece el mismo efecto diamagnético aun cuando un átomo ya tenga un momento permanente. Luego, el sistema precesará en el campo magnético. A medida que todo el átomo precesa, adquiere una pequeña velocidad angular adicional y ese girar lento da una pequeña corriente que representa una corrección al momento magnético. Éste es justamente el efecto diamagnético representado de otra manera. Sin embargo, realmente no tenemos que preocuparnos de eso al hablar del paramagnetismo. Si primero se calcula el efecto diamagnético, como hemos hecho aquí, no tenemos por qué preocuparnos de que haya una pequeña corriente adicional debida a la precesión. Esto ya ha sido incluido en el término diamagnético.

34-5 Teorema de Larmor

Ya podemos sacar una conclusión de los resultados obtenidos hasta ahora. Primero que todo, en la teoría clásica el momento $\boldsymbol{\mu}$ siempre era proporcional a $\mathbf{J}$ con una constante dada de proporcionalidad para un átomo en particular. No había ningún espín de los electrones y la constante de proporcionalidad era siempre $-q_e/2m$; es decir, en la ecuación (34.6) debíamos poner $g = 1$. El cociente $\boldsymbol{\mu}$ entre $\mathbf{J}$ era independiente de los movimientos internos de los electrones. Así, según la teoría clásica todos los sistemas de electrones precesarían con *la misma* velocidad angular (esto *no* es verdad en la mecánica cuántica). Este resultado está relacionado con un teorema de la mecánica clásica que desearíamos demostrar ahora. Supongan que tenemos un grupo de electrones que una atracción hacia un punto central mantiene juntos —como el núcleo atrae a los electrones—. Los electrones también estarán interactuando entre sí y, en general, pueden tener movimientos complicados. Además, supongan que han hallado la solución *sin* campo magnético y entonces quieren saber cuáles serían los movimientos *con* un campo magnético débil. El teorema dice que el movimiento en un campo magnético débil siempre es una de las soluciones sin campo más una rotación alrededor del eje del campo con una velocidad angular $\omega_L = q_e B/2m$. (Lo cual es igual a ω_p si $g = 1$.) Naturalmente, hay muchos movimientos posibles. Lo importante es que para cada movimiento sin campo magnético hay un movimiento correspondiente con el campo, el cual es el movimiento original más una rotación uniforme. Esto se llama teorema de Larmor y ω_L se llama *frecuencia de Larmor*.

Nos gustaría mostrar cómo se demuestra el teorema, pero les dejaremos calcular los detalles. Tomen primero un electrón en un campo de fuerzas centrales. La fuerza que experimenta es simplemente $\mathbf{F}(r)$ dirigida hacia el centro. Si ahora aplicamos un campo magnético uniforme, tenemos la fuerza adicional $q\mathbf{v} \times \mathbf{B}$; la fuerza total es entonces

$$\mathbf{F}(r) + q\mathbf{v} \times \mathbf{B}. \quad (34.18)$$

Examinemos ahora el mismo sistema desde un sistema de coordenadas que rota con velocidad angular ω alrededor de un eje paralelo a $\mathbf{B}$ que pasa por el centro de fuerza. Esto ya no es un sistema inercial, por lo que tenemos que introducir las pseudofuerzas apropiadas —las fuerzas centrífuga y de Coriolis de que hablamos en el capítulo 19 del volumen I—. Encontramos allí que en un sistema de ejes que rota con velocidad angular ω hay una fuerza *tangencial* aparente proporcional a v_r , la componente radial de la velocidad:

$$F_t = -2m\omega v_r. \quad (34.19)$$

Y hay una fuerza radial aparente que está dada por

$$F_r = m\omega^2 r + 2m\omega v_t, \quad (34.20)$$

donde v_t es la componente tangencial de la velocidad medida *en* el sistema rotante de coordenadas. (La componente radial v_r es la misma en los sistemas rotante e inercial.)

Ahora bien, para velocidades angulares suficientemente pequeñas (es decir, si $\omega r \ll v_t$), podemos despreciar el primer término (centrífugo) en la ecuación (34.20) frente al segundo (de Coriolis). Entonces es posible escribir conjuntamente las ecuaciones (34.19) y (34.20) en la forma

$$\mathbf{F} = -2m\boldsymbol{\omega} \times \mathbf{v}. \quad (34.21)$$

Si ahora *combinamos* una rotación y un campo magnético, debemos agregar la fuerza (34.21) a la fuerza (34.18). La fuerza total es

$$\mathbf{F}(r) + q\mathbf{v} \times \mathbf{B} + 2m\mathbf{v} \times \boldsymbol{\omega} \quad (34.22)$$

[invertimos el producto vectorial y el signo de la ecuación (34.21) para obtener el último término]. Examinando nuestro resultado, vemos que si

$$2m\boldsymbol{\omega} = -q\mathbf{B}$$

los dos últimos términos se cancelan y la única fuerza en el sistema rotante es $\mathbf{F}(r)$. El movimiento del electrón es exactamente lo mismo que sin campo magnético —y, por supuesto, sin rotación—. Hemos demostrado el teorema de Larmor para un electrón. Como la demostración supone una ω pequeña, el teorema es válido sólo para campos magnéticos débiles. Lo único que podríamos pedirles para mejorar las cosas es que tomen el caso de muchos electrones interactuando entre sí, pero todos en el mismo campo central, y que demuestren el mismo teorema. Así, no importa lo complicado que sea un átomo: si tiene un campo central el teorema es válido. Pero esto es el fin de la mecánica clásica, porque en realidad no es verdad que los movimientos precesen de esa manera. La frecuencia de precesión ω_p de la ecuación (34.11) sólo es igual a ω_L si ocurre que g es igual a 1.

34-6 La física clásica no da ni diamagnetismo ni paramagnetismo

Ahora nos gustaría demostrar que conforme a la mecánica clásica no puede haber ni diamagnetismo ni paramagnetismo. Parece una locura —primero demostramos que hay paramagnetismo, diamagnetismo, órbitas que precesan, etc.—, y ahora vamos a demostrar que todo está mal. ¡Sí! Vamos a demostrar que *si* siguen con la mecánica *clásica* suficientemente lejos, no hay tales efectos magnéticos —*todos se cancelan*—. Si empiezan un razonamiento clásico en cierto punto y no avanzan lo suficiente en las consecuencias, pueden obtener la respuesta que quieran. Pero la única demostración legítima y correcta muestra que no hay ningún efecto magnético.

Es una consecuencia de la mecánica clásica que si tienen cualquier clase de sistema —un gas con electrones, protones o lo que sea— dentro de una caja de modo que el todo no puede girar, no habrá efecto magnético. Es posible tener un efecto magnético si tienen un sistema aislado, como una estrella que se mantiene unida por sí misma, el cual puede comenzar a rotar cuando aplican el campo magnético. Pero si tienen un trozo de material que se mantiene en su sitio de modo que no puede empezar a girar, no habrá efectos magnéticos. Resumimos así lo que entendemos por impedir que gire: a una temperatura dada suponemos que hay *sólo un estado* de equilibrio termodinámico. El teorema dice entonces que si aplican un campo magnético y esperan que el sistema alcance equilibrio termodinámico, no habrá, ni paramagnetismo ni diamagnetismo —no habrá ningún momento magnético inducido—. Demostración: según la mecánica estadística, la probabilidad de que un sistema tenga un determinado estado de movimiento es proporcional a $e^{-U/kT}$, donde U es la energía de ese movimiento. Ahora bien: ¿cuál es la energía de movimiento? Para una partícula moviéndose en un campo magnético constante la energía es la energía potencial ordinaria más $mv^2/2$ sin nada adicional debido al campo magnético. [Ustedes saben que las fuerzas debidas a campos electromagnéticos son $q(\mathbf{E} + \mathbf{v} \times \mathbf{B})$, y que la rapidez $\mathbf{F} \cdot \mathbf{v}$ con que se realiza el trabajo es simplemente $q\mathbf{E} \cdot \mathbf{v}$, la cual no está afectada por el campo magnético.] Así, la energía de un sistema, esté o no en

un campo magnético, siempre está dada por la energía cinética más la energía potencial. Como la probabilidad de cualquier movimiento sólo depende de la energía —es decir, de la velocidad y de la posición— es lo mismo que haya o no un campo magnético. Por lo tanto, el campo magnético no tiene ningún efecto sobre el equilibrio *termodinámico*. Si tenemos un sistema en una caja y luego tenemos otro sistema en una segunda caja, esta vez con un campo magnético, la probabilidad de cualquier velocidad en cualquier punto de la primera caja es igual que en la segunda. Si la primera caja no tiene ninguna corriente media circulante (y no la tendrá si está en equilibrio con las paredes estáticas), no hay momento magnético medio. Como en la segunda caja todos los movimientos son iguales, tampoco allí hay un momento magnético medio. Por lo tanto, si se mantiene la temperatura constante y se restablece el equilibrio termodinámico después de aplicar el campo, no puede haber ningún momento magnético inducido por el campo —conforme a la mecánica clásica—. Sólo con la mecánica cuántica podemos llegar a una comprensión satisfactoria de los fenómenos magnéticos.

Desafortunadamente, no podemos suponer que ustedes tengan un buen conocimiento de la mecánica cuántica, por lo que difícilmente sea éste el momento de discutir el tema. Por otra parte, no siempre tenemos que aprender algo aprendiendo primero las reglas exactas y luego aprendiendo cómo se aplican en casos diferentes. Prácticamente todos los temas que hemos abordado en este curso han sido tratados de diferente modo. En el caso de la electricidad, escribimos las ecuaciones de Maxwell en la “página uno” y luego dedujimos todas las consecuencias. Ésa es una manera. Pero ahora *no* trataremos de empezar una nueva “página uno” escribiendo las ecuaciones de la mecánica cuántica y deduciendo todo de ellas. Simplemente les hablaremos de algunas consecuencias de la mecánica cuántica antes de que aprendan de dónde vienen. Así que, allá vamos.

34-7 Momentum angular en la mecánica cuántica

Ya les hemos dado una relación entre el momento magnético y el momentum angular. Eso está muy bien. Pero, ¿cuál es el *significado* del momento magnético y el momentum angular en la mecánica cuántica? En la mecánica cuántica resulta que lo mejor es definir cosas como los momentos magnéticos en términos de otros conceptos, como energía, para estar seguros de lo que significa. Ahora bien, es fácil definir un momento magnético en función de la energía porque la energía de un momento en un campo magnético es, en la teoría clásica, $\boldsymbol{\mu} \cdot \mathbf{B}$. En consecuencia, se ha tomado la siguiente definición en la mecánica cuántica: si calculamos la energía de un sistema en un campo magnético y encontramos que es proporcional a la intensidad de campo (para campos pequeños), el coeficiente se llama componente del momento magnético en la dirección del campo. (No necesitamos ponernos tan elegantes para nuestro trabajo; aún podemos pensar en el momento magnético en el sentido ordinario clásico, hasta cierto punto.)

Nos gustaría discutir ahora el concepto de momentum angular en la mecánica cuántica —o más bien las características de lo que en la mecánica cuántica se llama momentum angular—. Como ven, al pasar a nuevas clases de leyes, no pueden suponer simplemente que cada palabra va a significar exactamente lo mismo. Puede que piensen: “Ah, sé lo que es el momentum angular. Es eso que varía cuando aplicamos un torque”. ¿Pero qué es un torque? En mecánica cuántica tenemos que tener nuevas definiciones de las viejas magnitudes. Por lo tanto, legalmente, lo mejor sería darle otro nombre, como “momentum cuantangular” o algo por el estilo, porque es el momentum angular como se lo define en la mecánica cuántica. Pero si en la mecánica cuántica podemos hallar una magnitud que sea idéntica a nuestra vieja idea de momentum angular cuando el sistema se hace suficientemente grande, no es de utilidad inventar una palabra más. También podríamos llamarlo momentum angular simplemente. Aclarado esto, la cosa rara que estamos a punto de describir *es* el momentum angular. Es lo que en sistemas grandes reconocemos como momentum angular en la mecánica clásica.

Primero, tomemos un sistema cuyo momentum angular se conserva, como un átomo solitario en el espacio vacío. Ese sistema podría (como la Tierra girando sobre su eje) estar girando, en el sentido ordinario de la palabra, alrededor de cualquier eje que se elija. Y para cada velocidad de giro podría haber muchos “estados” diferentes, todos de la misma energía correspondiendo cada “estado” a una dirección determinada del eje del momentum angular. Así, en la teoría clásica, dado un momentum angular, hay un número infinito de estados posibles, todos con la misma energía.

Por el contrario, en la mecánica cuántica resulta que ocurren varias cosas extrañas. Primero el número de estados en que ese sistema *puede existir* está limitado —sólo hay un número finito—. Si el sistema es pequeño, el número finito es muy pequeño y si el sistema es grande, el número finito se hace muy, pero muy grande. Segundo: *no podemos* describir un “estado” dando la *dirección* angular según alguna dirección —la dirección z , por ejemplo—. Clásicamente, un objeto con un momentum angular total J podría tener, para su componente z , cualquier valor entre $+J$ y $-J$. Pero cuánticamente, la componente z del momentum angular sólo puede tener ciertos valores discretos. Cualquier sistema —un átomo, o un núcleo, o lo que sea— con cierta energía dada, tiene un número característico j y la componente z de su momentum angular sólo puede asumir uno de los siguientes valores:

$$\begin{aligned} & j\hbar \\ & (j-1)\hbar \\ & (j-2)\hbar \\ & \vdots \\ & -(j-2)\hbar \\ & -(j-1)\hbar \\ & -j\hbar \end{aligned} \tag{34.23}$$

La componente z más grande es j por $\hbar$; la que le sigue es una unidad de $\hbar$ menor, y así siguiendo hasta $-j\hbar$. El número j se llama “espín del sistema” (algunos lo llaman “número cuántico del momentum angular total”, pero nosotros lo llamaremos “espín”).

Puede que les preocupe saber si lo que estamos diciendo vale únicamente para cierto eje z “especial”. Pero no es así. Para un sistema cuyo espín es j , la componente del momentum angular según *cualquier* eje sólo puede tener uno de los valores en (34.23). Aunque es bastante misterioso, les pedimos que lo acepten por el momento. Más adelante volveremos sobre este punto y lo discutiremos. Por lo menos pueden estar contentos de oír que la componente z va desde cierto número hasta menos el *mismo* número, para que por lo menos no tengamos que decidir cuál es la dirección positiva del eje z . (Por cierto que si dijéramos que va desde $+j$ hasta menos una cantidad diferente, eso sería infinitamente misterioso porque no podríamos definir el eje z apuntando en el otro sentido.)

Ahora bien, si la componente z del momentum angular debe bajar de uno en uno desde $+j$ hasta $-j$, entonces j debe ser un entero. ¡No! No tanto; dos veces j debe ser un entero. Sólo la *diferencia* entre $+j$ y $-j$ debe ser un entero. Así, en general, el espín j es entero o semientero, según que $2j$ sea par o impar. Por ejemplo, tomen un núcleo como el litio, que tiene espín tres medios: $j = 3/2$. Entonces el momentum angular respecto al eje z es, en unidades de $\hbar$, uno de los siguientes:

$$\begin{aligned} & +3/2 \\ & +1/2 \\ & -1/2 \\ & -3/2. \end{aligned}$$

Hay cuatro estados posibles, todos de la misma energía, si el núcleo está en el espacio vacío sin campos externos. Si tenemos un sistema cuyo espín es dos,

la componente z de su momentum angular tiene únicamente los siguientes valores, en unidades de $\hbar$,

$$\begin{array}{c} 2 \\ 1 \\ 0 \\ -1 \\ -2. \end{array}$$

Si cuentan cuántos estados hay para un j dado, encuentran $(2j+1)$ posibilidades. En otras palabras, si me dicen cuál es la energía y también el espín j , resulta que hay exactamente $(2j+1)$ estados con esa energía, cada uno de los cuales corresponde a uno de los diferentes valores posibles de la componente z del momentum angular.

Agreguemos otro hecho. Si toman al azar cualquier átomo de j conocido y miden la componente z del momentum angular, entonces pueden obtener cualquiera de los valores posibles, y cada uno de ellos es *igualmente* probable. En efecto, todos los estados son únicos y cada uno es tan bueno como el otro. Cada uno tiene el mismo “peso” en el mundo. (Estamos suponiendo que no se ha hecho nada para escoger una muestra especial.) Incidentalmente, este hecho tiene un análogo clásico simple, si hacen la misma pregunta clásicamente: ¿cuál es la probabilidad de un valor particular de la componente z del momentum angular si toman una muestra al azar de sistemas, todos con el mismo momentum angular total? —la respuesta es que todos los valores desde el máximo hasta el mínimo son igualmente probables (pueden calcular esto fácilmente)—. El resultado clásico corresponde a la equiprobabilidad de las $(2j+1)$ posibilidades en la mecánica cuántica.

Con lo que tenemos hasta ahora podemos obtener otra conclusión interesante y algo sorprendente. En ciertos cálculos clásicos, la cantidad que aparece en el resultado final es el *cuadrado* del módulo del momentum angular $\mathbf{J}$ —o sea $\mathbf{J} \cdot \mathbf{J}$ —. Resulta que muchas veces es posible *adivinar* la fórmula cuántica correcta usando el cálculo clásico y la siguiente regla simple: reemplazar $J^2 = \mathbf{J} \cdot \mathbf{J}$ por $j(j+1)\hbar^2$. Esta regla se usa mucho y usualmente da el resultado correcto, pero *no* siempre. Podemos dar el siguiente argumento para mostrar por qué sería de esperar que esta regla funcione.

Se puede escribir el producto escalar $\mathbf{J} \cdot \mathbf{J}$ en la forma

$$\mathbf{J} \cdot \mathbf{J} = J_x^2 + J_y^2 + J_z^2.$$

Como es un escalar, debe ser el mismo para cualquier orientación del espín. Supongan que tomamos muestras al azar de un sistema determinado cualquiera y hacemos medidas de J_x^2 , J_y^2 o J_z^2 , el *valor medio* debe ser igual para todas (ninguna dirección es privilegiada). En consecuencia, el promedio de $\mathbf{J} \cdot \mathbf{J}$ es simplemente igual a tres veces el promedio de cualquiera de las componentes al cuadrado, digamos que de J_z^2 :

$$\langle \mathbf{J} \cdot \mathbf{J} \rangle_{\text{prom}} = 3\langle J_z^2 \rangle_{\text{prom}}.$$

Pero como $\mathbf{J} \cdot \mathbf{J}$ es el mismo para todas las orientaciones, su promedio es, naturalmente, su valor constante; tenemos

$$\mathbf{J} \cdot \mathbf{J} = 3\langle J_z^2 \rangle_{\text{prom}}. \quad (34.24)$$

Si ahora decimos que vamos a usar la misma ecuación en la mecánica cuántica, podemos hallar fácilmente $\langle J_z^2 \rangle_{\text{prom}}$. Sólo tenemos que sumar los $(2j+1)$ valores posibles de J_z^2 y dividir entre el número total:

$$\langle J_z^2 \rangle_{\text{prom}} = \frac{j^2 + (j-1)^2 + \cdots + (-j+1)^2 + (-j)^2}{2j+1} \hbar^2. \quad (34.25)$$

Para un sistema de espín $3/2$ se tiene:

$$\langle J_z^2 \rangle_{\text{prom}} = \frac{(3/2)^2 + (1/2)^2 + (-1/2)^2 + (-3/2)^2}{4} \hbar^2 = \frac{5}{4} \hbar^2.$$

Concluimos que

$$\mathbf{J} \cdot \mathbf{J} = 3\langle J_z^2 \rangle_{\text{prom}} = 3\frac{5}{4}\hbar^2 = \frac{3}{2}\left(\frac{3}{2} + 1\right)\hbar^2.$$

Dejamos que ustedes demuestren que la ecuación (34.25), junto con la ecuación (34.24), da el resultado general

$$\mathbf{J} \cdot \mathbf{J} = j(j+1)\hbar^2. \quad (34.26)$$

Aunque clásicamente pensaríamos que el mayor valor posible de la componente z de $\mathbf{J}$ es justamente el módulo de $\mathbf{J}$ —dicho de otro modo, $\sqrt{\mathbf{J} \cdot \mathbf{J}}$ —cuánticamente, el máximo de J_z siempre es un poco menor porque $j\hbar$ siempre es menor que $\sqrt{j(j+1)}\hbar$. Además, el momentum angular nunca yace “completamente en la dirección z ”.

34-8 Energía magnética de los átomos

Ahora queremos hablarles nuevamente del momento magnético. Hemos dicho que en la mecánica cuántica el momento magnético de un sistema atómico determinado se puede escribir en función del momentum angular mediante la ecuación (34.6),

$$\boldsymbol{\mu} = -g\left(\frac{q_e}{2m}\right)\mathbf{J}, \quad (34.27)$$

donde $-q_e$ y m son la carga y la masa del electrón.

Un imán atómico colocado en un campo magnético externo tendrá una energía magnética adicional que depende de la componente de su momento magnético en la dirección del campo. Sabemos que

$$U_{\text{mag}} = -\boldsymbol{\mu} \cdot \mathbf{B}. \quad (34.28)$$

Eligiendo el eje z en la dirección de $\mathbf{B}$,

$$U_{\text{mag}} = -\mu_z B. \quad (34.29)$$

Usando la ecuación (34.27) tenemos

$$U_{\text{mag}} = g\left(\frac{q_e}{2m}\right)J_z B.$$

La mecánica cuántica afirma que J_z solamente puede tener ciertos valores: $j\hbar$, $(j-1)\hbar$, $\dots$, $-j\hbar$. En consecuencia, la energía magnética de un sistema atómico no es arbitraria; sólo puede tener ciertos valores. Por ejemplo, su máximo es

$$g\left(\frac{q_e}{2m}\right)\hbar j B.$$

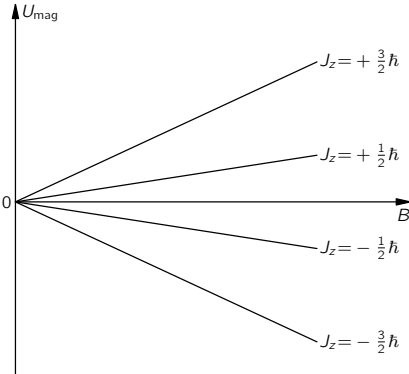


Figura 34-5. Energías magnéticas posibles de un sistema atómico de espín $3/2$ en un campo magnético $\mathbf{B}$.

La cantidad $q_e\hbar/2m$ recibe habitualmente el nombre de “magnetón de Bohr” y se escribe μ_B :

$$\mu_B = \frac{q_e\hbar}{2m}.$$

Los valores posibles de la energía magnética son

$$U_{\text{mag}} = g\mu_B B \frac{J_z}{\hbar},$$

donde $J_z/\hbar$ asume los valores posibles j , $(j-1)$, $(j-2)$, $\dots$, $(-j+1)$, $-j$.

En otras palabras, cuando se pone un sistema atómico en un campo magnético su energía varía en una cantidad que es proporcional al campo y a J_z . Decimos que la energía de un sistema atómico “se desdobra en $2j+1$ niveles” debido al campo magnético. Por ejemplo, un átomo cuya energía es U_0 fuera de un campo magnético y cuyo j es $3/2$, tendrá cuatro energías posibles al colocarlo en un campo. Podemos mostrarles estas energías por medio de un diagrama de niveles

de energía como el dibujado en la figura 34-5. Cualquier átomo en particular sólo puede tener una de las cuatro energías posibles en un campo determinado B . Esto es lo que la mecánica cuántica dice sobre el comportamiento de un sistema atómico en un campo magnético.

El sistema “atómico” más simple es un electrón solo. El espín de un electrón es $1/2$, por lo que hay dos estados posibles: $J_z = \hbar/2$ y $J_z = -\hbar/2$. Para un electrón en reposo (sin movimiento orbital) el momento magnético de espín tiene un g que vale 2, de modo que la energía magnética puede ser $\pm\mu_B B$. La figura 34-6 muestra las energías posibles en un campo magnético. Hablando un poco a la ligera decimos que el electrón tiene su espín o bien, “hacia arriba” (según el campo), o bien “hacia abajo” (opuesto al campo).

Para sistemas de espín más alto hay más estados. Podemos pensar que el espín está “hacia arriba” o “hacia abajo” o inclinado a cierto “ángulo” intermedio, según sea el valor de J_z .

Usaremos estos resultados cuánticos para estudiar las propiedades magnéticas de los materiales en el próximo capítulo.

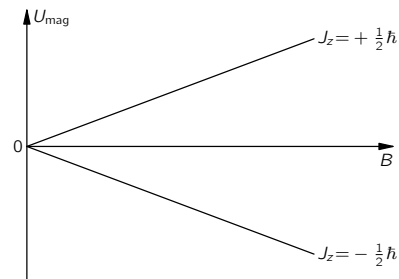


Figura 34-6. Los dos estados posibles de energía de un electrón en un campo magnético B .

Paramagnetismo y resonancias magnéticas

35-1 Estados magnéticos cuantizados

En el último capítulo describimos cómo es que en la mecánica cuántica el momentum angular de algo no tiene una dirección arbitraria, sino que su componente según un eje dado sólo puede tomar ciertos valores discretos igualmente espaciados. Es algo raro y peculiar. Quizá piensen que no deberíamos entrar en tales cosas hasta que ustedes estén más avanzados y listos para aceptar esta clase de idea. En realidad, sus mentes nunca estarán suficientemente listas —en el sentido de que sean capaces de aceptar fácilmente tal cosa—. No hay modo descriptivo alguno de hacerla inteligible y que a la vez no sea tan sutil y avanzado en su propia forma que resulte más complicado que lo que tratan de explicar. El comportamiento de la materia a escala microscópica es verdaderamente muy extraño —como hemos señalado muchas veces— y es diferente de cualquier cosa a la que están acostumbrados. A medida que avanzamos en la física clásica, es una buena idea tratar de irse familiarizando poco a poco con el comportamiento de las cosas a una escala microscópica, al principio como una experiencia sin ninguna comprensión profunda. La comprensión de ello viene muy lentamente, si es que viene. Naturalmente, uno es capaz de saber qué va a suceder en una situación cuántica —si esto es lo que comprensión significa—, pero nunca se tiene la sensación de comodidad de que estas reglas cuánticas son “naturales”. Por supuesto que lo *son*, pero no son naturales para nuestra experiencia propia a un nivel ordinario. Debemos aclarar que la actitud que vamos a tomar respecto a esta regla del momentum angular es completamente diferente de muchas de las otras cuestiones de que hemos hablado. No vamos a tratar de “explicarla”, pero al menos debemos *decirles* lo que sucede; no sería honrado describirles las propiedades magnéticas de los materiales sin mencionarles el hecho de que la descripción clásica del magnetismo —del momentum angular y de los momentos magnéticos— es incorrecta.

Una de las características más sorprendentes y perturbadoras de la mecánica cuántica es que si toman el momentum angular según cualquier eje particular encuentran que siempre es un entero o un semientero por $\hbar$. Esto es así, sea cual sea el eje que tomen. Las sutilezas de este hecho curioso —que pueden tomar cualquier otro eje y encuentran que la componente según él también está incluida en el mismo conjunto de valores— las dejaremos para un capítulo posterior, cuando experimentarán la satisfacción de ver cómo se resuelve finalmente esta aparente paradoja.

Ahora aceptaremos solamente el hecho de que para cada sistema atómico hay un número j , llamado *espín* del sistema —que debe ser un entero o un semientero— y que la componente del momentum angular según cualquier eje particular tendrá siempre uno de los siguientes valores entre $+j\hbar$ y $-j\hbar$:

$$J_z = \text{uno de } \begin{pmatrix} j \\ j-1 \\ j-2 \\ \vdots \\ -j+2 \\ -j+1 \\ -j \end{pmatrix} \cdot \hbar. \quad (35.1)$$

También hemos mencionado que todo sistema atómico sencillo posee un momento magnético que posee la misma dirección que el momentum angular.

- 35-1 Estados magnéticos cuantizados
- 35-2 El experimento de Stern-Gerlach
- 35-3 El método de haces moleculares de Rabi
- 35-4 Paramagnetismo macroscópico de materiales
- 35-5 Enfriamiento por desmagnetización adiabática
- 35-6 Resonancia magnética nuclear

Revisar

Volumen II, capítulo 11,
Dentro de los dieléctricos

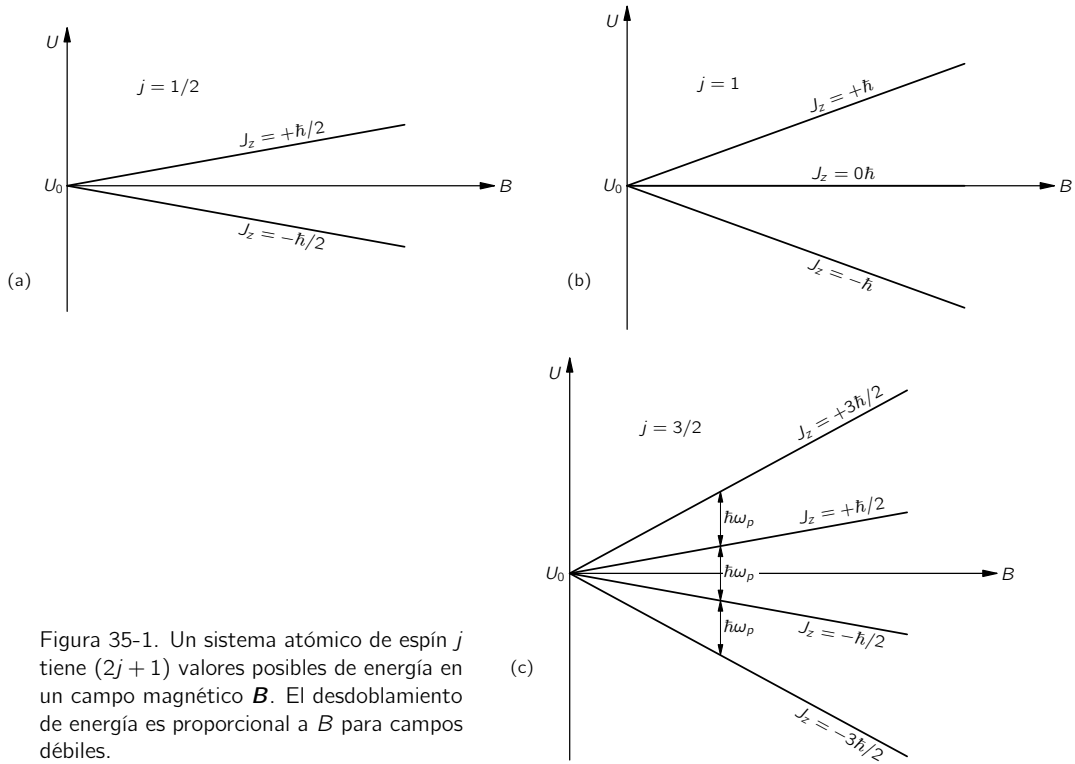


Figura 35-1. Un sistema atómico de espín j tiene $(2j + 1)$ valores posibles de energía en un campo magnético B . El desdoblamiento de energía es proporcional a B para campos débiles.

Esto no sólo es válido para átomos y núcleos sino también para las partículas fundamentales. Toda partícula fundamental tiene su valor propio característico para j y para el momento magnético. (Para algunas partículas ambos son cero.) Lo que queremos decir por “el momento magnético” en esa afirmación es que la energía del sistema en un campo magnético, en la dirección z por ejemplo, se puede escribir como $-\mu_z B$ para campos magnéticos débiles. Debemos imponer la condición de que el campo no debe ser muy intenso, porque de otro modo podría perturbar los movimientos internos del sistema y la energía no sería una medida del momento magnético que se encontraba presente antes de que se aplicase el campo. Pero si el campo es suficientemente débil, el mismo varía la energía en la cantidad

$$\Delta U = -\mu_z B, \quad (35.2)$$

siendo que en esta ecuación tenemos que remplazar μ_z por

$$\mu_z = g \left(\frac{q}{2m} \right) J_z, \quad (35.3)$$

donde J_z tiene uno de los valores de la ecuación (35.1).

Supongan que tenemos un sistema de espín $j = 3/2$. En ausencia de un campo magnético, el sistema tiene cuatro estados diferentes posibles correspondientes a los distintos valores de J_z , todos los cuales tienen exactamente la misma energía. Sin embargo, el momento en que introducimos el campo magnético, hay una energía adicional de interacción que separa estos estados en cuatro niveles de energía levemente diferentes. Las energías de estos niveles están dadas por una cierta energía proporcional a B , multiplicada por $\hbar$ por $3/2$, $1/2$, $-1/2$ y $-3/2$ —los valores de J_z —. Los diagramas de la figura 35-1 muestran los desdoblamientos de los niveles de energía para sistemas atómicos con espines $1/2$, 1 y $3/2$ (recuerden que el momento magnético siempre es antiparalelo al momentum angular para cualquier ordenamiento de electrones).

Observarán en los diagramas que el “centro de gravedad” de los niveles de energía es el mismo con o sin campo magnético. Observen también que los

espaciamientos de un nivel al siguiente son siempre iguales para una partícula dada en un campo magnético dado. Vamos a escribir el espaciamiento de energía, para un campo magnético B dado, como $\hbar\omega_p$ —lo cual es simplemente una definición de ω_p —. Usando las ecuaciones (35.2) y (35.3), tenemos

$$\hbar\omega_p = g \frac{q}{2m} \hbar B$$

o sea

$$\omega_p = g \frac{q}{2m} B. \quad (35.4)$$

La cantidad $g(q/2m)$ es justamente el cociente entre el momento magnético y el momentum angular —es una propiedad de la partícula—. La ecuación (35.4) es la misma fórmula que obtuvimos en el capítulo 34 para la velocidad angular de precesión en un campo magnético de un giroscopio cuyo momentum angular es $\mathbf{J}$ y cuyo momento magnético es $\boldsymbol{\mu}$.

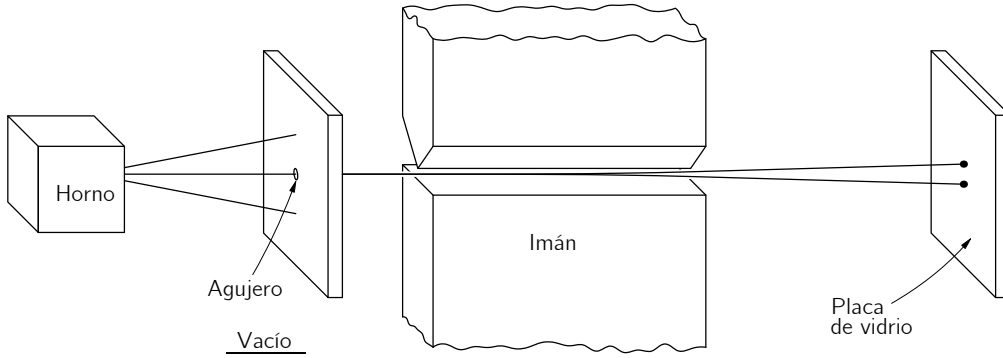


Figura 35-2. El experimento de Stern y Gerlach.

35-2 El experimento de Stern-Gerlach

El hecho de que el momentum angular esté cuantizado es tan sorprendente que hablaremos un poco de él históricamente. Su descubrimiento produjo una gran sorpresa (aunque teóricamente se esperaba). Stern y Gerlach lo observaron por primera vez en un experimento realizado en 1922. Si lo desean, pueden considerar el experimento de Stern-Gerlach como una justificación directa de la creencia de la cuantización del momentum angular. Stern y Gerlach idearon un experimento para medir el momento magnético de un átomo de plata. Produjeron un haz de átomos de plata evaporando plata en un horno y dejando que algunos salieran a través de una serie de agujeros pequeños. Este haz se dirigió entre los polos de un imán especial, como muestra la figura 35-2. Su idea fue la siguiente: si el átomo de plata tiene un momento magnético $\boldsymbol{\mu}$ y lo colocamos en un campo magnético $\mathbf{B}$, tendrá una energía $-\mu_z B$, siendo z la dirección del campo magnético. En la teoría clásica μ_z sería igual al momento magnético por el coseno del ángulo formado por el momento y el campo magnético, por lo que la energía adicional en el campo sería

$$\Delta U = -\mu B \cos \theta. \quad (35.5)$$

Naturalmente, al salir los átomos del horno sus momentos magnéticos apuntarían cualquier dirección posible y encontraríamos cualquier valor para θ . Ahora bien, si el campo magnético varía muy rápidamente con z —si hay un gradiente intenso del campo— entonces la energía magnética variará también con la posición y habrá una fuerza sobre los momentos magnéticos cuya dirección dependerá de si el coseno de θ es positivo o negativo. Los átomos serán atraídos hacia arriba o

hacia abajo de una fuerza proporcional a la derivada de la energía magnética; según el principio de trabajos virtuales, tenemos

$$F_z = -\frac{\partial U}{\partial z} = \mu \cos \theta \frac{\partial B}{\partial z}. \quad (35.6)$$

Stern y Gerlach construyeron su imán con uno de los polos muy en punta para que así se produjese una variación muy rápida del campo magnético. El haz de átomos de plata se dirigió justamente a lo largo de este filo para que así los átomos se viesan sometidos a una fuerza vertical en un campo no homogéneo. Un átomo de plata con su momento magnético dirigido horizontalmente no experimentaría fuerza y pasaría en línea recta el imán. Un átomo cuyo momento magnético estuviese exactamente vertical experimentaría una fuerza vertical que lo empujaría hacia el polo en la punta del imán. Un átomo cuyo momento magnético apuntase hacia abajo experimentaría un empuje hacia abajo. Por consiguiente, cuando abandonasen el imán, los átomos se distribuirían de acuerdo a sus componentes verticales de momento magnético. Todos los ángulos son posibles en la teoría clásica, de modo que cuando los átomos de plata se recogen por deposición en una placa de vidrio, se espera una distribución de plata según una línea vertical. La altura de la línea sería proporcional al módulo del momento magnético. El tremendo fracaso de las ideas clásicas se puso de manifiesto cuando Stern y Gerlach vieron lo que en la realidad sucede. Encontraron en la lámina de vidrio dos manchas nítidas. Los átomos de plata habían formado dos haces.

Es casi milagroso que un haz de átomos cuyos espines estarían aparentemente orientados al azar se bifurque en dos haces separados. ¿Cómo *sabe* el momento magnético que solamente se le permite llevar ciertas componentes en la dirección del campo magnético? En realidad éste fue el comienzo del descubrimiento de la cuantización del momentum angular, y en lugar de tratar de darles una explicación teórica, solamente les diré que se atengan al resultado del experimento, como los físicos de entonces tuvieron que aceptarlo en su momento cuando se realizó el experimento por primera vez. Es un *hecho experimental* que la energía de un átomo en un campo magnético toma una serie de valores particulares. Para cada uno de estos valores la energía es proporcional a la intensidad del campo. Por consiguiente, en una región donde el campo varía, el principio de los trabajos virtuales nos dice que la fuerza magnética posible sobre los átomos tendrá un conjunto de valores discretos; la fuerza es diferente para cada estado y, por lo tanto, el haz de átomos se separa en un número pequeño de haces distintos. Se puede encontrar la intensidad del momento magnético a partir de la medida de la desviación de los haces.

35-3 El método de haces moleculares de Rabi

Describiremos ahora un aparato mejor que mide momentos magnéticos y que fue desarrollado por Isidor Isaac Rabi y sus colaboradores. La desviación de los átomos en el experimento de Stern-Gerlach es muy pequeña y la medida del momento magnético no es muy precisa. La técnica de Rabi permite una precisión fantástica en la medida de los momentos magnéticos. El método se basa en el hecho de que la energía original de los átomos en un campo magnético se separa en un número finito de niveles de energía. Que la energía de un átomo en un campo magnético pueda tener solamente ciertas energías discretas no es realmente más sorprendente, que el hecho de que los átomos *en general* puedan tener solamente ciertos niveles discretos de energía —algo que mencionamos con frecuencia en el volumen I—. ¿Por qué *no* valdría lo mismo para átomos en un campo magnético? Sí, sirve. Pero se intenta correlacionarlo con la idea de un *momento magnético orientado* que implica algunas complicaciones raras de la mecánica cuántica.

Cuando un átomo tiene dos niveles que difieren en energía en la cantidad ΔU , puede realizar una transición del nivel superior al nivel inferior mediante la emisión de un cuanto de luz de frecuencia ω , siendo

$$\hbar\omega = \Delta U. \quad (35.7)$$

Lo mismo puede suceder con átomos en un campo magnético. Sólo que las diferencias de energía son tan pequeñas que la frecuencia no corresponde a luz sino a microondas o a radiofrecuencias. Las transiciones del nivel inferior de energía a un nivel superior de un átomo, también pueden tener lugar con la ayuda de la absorción de luz o, en el caso de átomos en un campo magnético, mediante la absorción de energía de microondas. Así, si tenemos un átomo en un campo magnético, podemos causar transiciones de un estado a otro mediante la aplicación de un campo electromagnético adicional con frecuencia apropiada. En otras palabras, si tenemos un átomo en un campo magnético intenso y le “hacemos cosquillas” al átomo con campo electromagnético débil que varía, habrá una cierta probabilidad de pasarlo a otro nivel si la frecuencia es próxima a la ω de la ecuación (35.7). Para un átomo en un campo magnético, esta frecuencia es justamente lo que antes llamamos ω_p y está dada en función del campo magnético por la ecuación (35.4). Si se cosquillea al átomo con la frecuencia incorrecta, la probabilidad de causar una transición es muy pequeña. Así, en la probabilidad de causar una transición hay una *resonancia* aguda en ω_p . Midiendo la frecuencia de esta resonancia en un campo magnético B conocido, podemos medir la cantidad $g(q/2m)$ —y por tanto el factor g — con gran precisión.

Es interesante cuando se llega a la misma conclusión desde un punto de vista clásico. Según la imagen clásica, cuando colocamos un pequeño giroscopio de momento magnético μ y momentum angular J en un campo magnético externo, el giroscopio precesará alrededor de un eje paralelo al campo magnético (véase figura 35-3). Supongan que preguntamos: ¿cómo podemos cambiar el ángulo del giroscopio clásico respecto al campo —a saber, respecto al eje z —? El campo magnético produce un torque en dirección *horizontal*. Pensarían que tal torque está *tratando* de alinear el imán con el campo, pero solamente causa la precesión. Si queremos cambiar el ángulo del giroscopio respecto al eje z , debemos aplicarle un torque *en dirección del eje z* . Si aplicamos un torque que va en la misma dirección que la precesión, el ángulo del giroscopio cambiará y dará una componente más pequeña de J en la dirección z . En la figura 35-3, aumentará el ángulo entre J y el eje z . Si tratamos de impedir la precesión, J se mueve sobre la vertical.

¿Cómo podemos aplicar la clase de torque que deseamos a nuestro átomo precesando en un campo magnético uniforme? La respuesta es: aplicando un pequeño campo magnético lateral. Puede que piensen que la dirección de este campo magnético tendría que rotar con la precesión del momento magnético de manera que siempre formase con él un ángulo recto, como el campo B' en la figura 35-4(a). Este campo sirve, pero un campo horizontal *alterno* es casi tan bueno. Si tenemos un campo horizontal B' pequeño, el cual está siempre en la dirección x (más o menos) y que oscila con la frecuencia ω_p en cada mitad de ciclo el torque sobre el momento magnético se invierte, de modo que tiene un efecto acumulativo que es casi tan efectivo como un campo magnético rotatorio. Clásicamente, entonces, esperaríamos que la componente del momento magnético según la dirección z cambiase si tenemos un campo magnético oscilatorio muy débil con una frecuencia exactamente ω_p . Es natural que, clásicamente, μ_z cambiaría continuamente, pero en la mecánica cuántica la componente z del momento magnético no se puede ajustar continuamente. Debe saltar de repente de un valor a otro. Hemos hecho la comparación entre las consecuencias de la mecánica clásica y de la mecánica cuántica para darles cierta idea de lo que puede suceder clásicamente y de cómo se relaciona con lo que realmente sucede en la mecánica cuántica. A propósito, notarán que la frecuencia de resonancia esperada es la misma en ambos casos.

Una nota adicional: según lo que hemos dicho acerca de la mecánica cuántica, no hay razón aparente para que no haya también transiciones a la frecuencia $2\omega_p$. Resulta que no hay nada análogo en el caso clásico, ni tampoco sucede en la teoría cuántica —por lo menos en el método particular de inducir las transiciones que hemos descrito—. Con un campo magnético horizontal oscilante, la probabilidad de que una frecuencia $2\omega_p$ cause de un salto de dos pasos al mismo tiempo es cero.

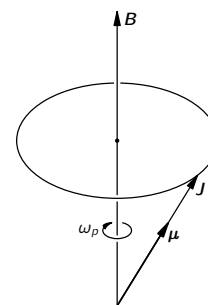


Figura 35-3. La precesión clásica de un átomo con momento magnético μ y momento angular J .

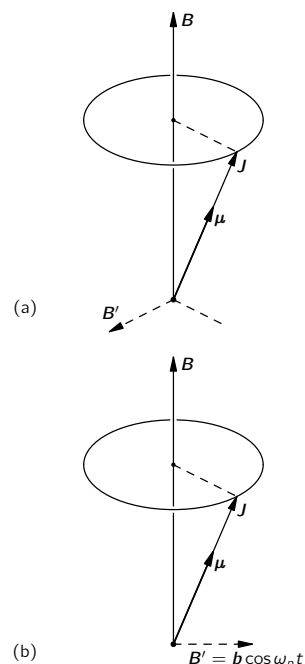


Figura 35-4. El ángulo de precesión de un imán atómico se puede variar por medio de un campo magnético horizontal siempre perpendicular a μ , como en (a), o mediante un campo oscilante, como en (b).

Si el campo oscilante $\mathbf{B}'$ tiene una frecuencia apreciablemente diferente de ω_p , no invertirá espines y los átomos seguirán sus caminos no perturbados hacia el detector. De este modo pueden ver que la frecuencia de “precesión” ω_p de los átomos en el campo $\mathbf{B}_0$ se puede encontrar variando la frecuencia ω del campo $\mathbf{B}'$ hasta que se observe una disminución en la corriente de átomos que llega al detector. Ocurrirá una disminución de la corriente cuando ω esté “en resonancia” con ω_p . Una representación de la corriente del detector en función de ω se puede parecer a lo que muestra la figura 35-6. Conociendo ω_p podemos obtener el valor de g del átomo.

Estos experimentos con haces atómicos, o como se les llama comúnmente, experimentos de resonancia de haces “moleculares”, son una manera delicada y bella de medir las propiedades magnéticas de los objetos atómicos. Se puede determinar con gran precisión la frecuencia de resonancia ω_p —de hecho, con una precisión mayor que con la que medimos el campo magnético $\mathbf{B}_0$, el cual debemos conocer para encontrar g —.

35-4 Paramagnetismo macroscópico de materiales

Ahora desearíamos describir el fenómeno de paramagnetismo macroscópico de los materiales. Supongan que tenemos una sustancia cuyos átomos tienen momentos magnéticos permanentes, por ejemplo, un cristal de sulfato de cobre. En el cristal hay iones de cobre cuyas capas electrónicas internas tienen un momentum angular resultante y un momento magnético resultante. Así, el ion cobre es un objeto que tiene un momento magnético permanente. Digamos unas palabras sobre cuáles átomos tienen momentos magnéticos y cuáles no. Cualquier átomo, el sodio por ejemplo, que tenga un número *impar* de electrones, tendrá momento magnético. El sodio tiene un electrón en su capa incompleta y este electrón da al átomo un espín y un momento magnético. Sin embargo, comúnmente cuando se forman los compuestos, los electrones adicionales de la capa externa se acoplan con otros electrones cuyos espines tienen direcciones opuestas, de modo que todos los momenta angulares y todos los momentos magnéticos de los electrones de valencia se cancelan. Ésta es la razón, en general, para que las moléculas no tengan momento magnético. Por supuesto, si tienen un gas de átomos de sodio, no existe la compensación.* Del mismo modo, si tienen lo que se llama en química un “radical libre” —un objeto con un número impar de electrones de valencia— los enlaces no se saturan completamente y hay un momentum angular resultante.

En la mayoría de los materiales macroscópicos hay un momento magnético resultante solamente si hay átomos presentes cuya capa electrónica *interna* no está completa. En este caso, puede haber un momentum angular resultante y un momento magnético. Estos átomos se encuentran en la parte de los “elementos de transición” de la tabla periódica —por ejemplo, el cromo, manganeso, hierro, níquel, cobalto, paladio y el platino son elementos de esta clase—. Todos los elementos de las tierras raras también tienen capas internas incompletas y, por lo tanto, tienen momentos magnéticos permanentes. Hay un par de elementos extraños que también tienen momentos magnéticos; uno de ellos es el oxígeno líquido, pero dejaremos que el departamento de química les explique la razón.

Supongan ahora que tenemos una caja llena de átomos o moléculas con momentos permanentes —digamos un gas, un líquido o un cristal—. Nos gustaría saber qué sucede si aplicamos un campo magnético externo. Sin *ningún* campo magnético, los átomos son empujados desordenadamente por los movimientos térmicos y los momentos terminan apuntando en todas direcciones. Pero cuando hay un campo magnético, éste trata de alinear los imanes pequeños; hay entonces más momentos en la dirección del campo que en otra dirección. El material se ha “magnetizado”.

Definimos la *magnetización* $\mathbf{M}$ de un material como el momento magnético resultante por unidad de volumen, con lo cual entendemos la suma vectorial de

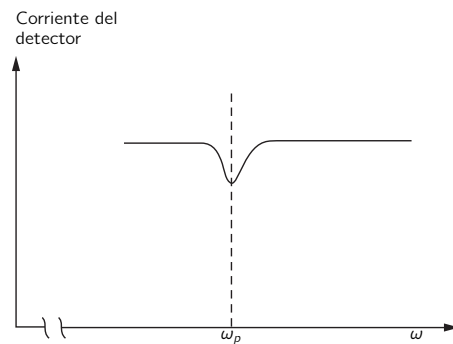


Figura 35-6. La corriente de átomos del haz disminuye cuando $\omega = \omega_p$.

* Generalmente el vapor de Na es en su mayoría monoatómico, a menos que haya algunas moléculas de Na_2 .

todos los momentos magnéticos atómicos en una unidad de volumen. Si hay N átomos por unidad de volumen y su momento *promedio* es $\langle \boldsymbol{\mu} \rangle_{\text{prom}}$, entonces $\mathbf{M}$ se puede expresar como N por el momento atómico promedio:

$$\mathbf{M} = N \langle \boldsymbol{\mu} \rangle_{\text{prom}}. \quad (35.8)$$

La definición de $\mathbf{M}$ corresponde a la definición de polarización eléctrica $\mathbf{P}$ del capítulo 10.

La teoría clásica del paramagnetismo es exactamente como la teoría de la constante dieléctrica que dimos en el capítulo 11. Se supone que cada uno de los átomos tenga un momento magnético $\boldsymbol{\mu}$, que siempre tiene el mismo módulo, pero que puede apuntar en cualquier dirección. En un campo $\mathbf{B}$, la energía magnética es $-\boldsymbol{\mu} \cdot \mathbf{B} = -\mu B \cos \theta$, donde θ es el ángulo formado por el momento y el campo. Según la mecánica estadística, la probabilidad relativa de tener cualquier ángulo es $e^{-\text{energía}/kT}$ por lo que los ángulos próximos a cero son más probables que los próximos a π . Procediendo exactamente como hicimos en la sección 11-3, encontramos que para campos magnéticos pequeños, $\mathbf{M}$ es paralelo a $\mathbf{B}$ y su módulo es

$$M = \frac{N\mu^2 B}{3kT}. \quad (35.9)$$

[Véase ecuación (11.20).] Esta fórmula aproximada es correcta solamente para $\mu B/kT$ mucho menor que uno.

Encontramos que la magnetización inducida —el momento magnético por unidad de volumen— es proporcional al campo magnético. Éste es el fenómeno del paramagnetismo. Verán que el efecto es más intenso a bajas temperaturas y más débil a temperaturas más altas. Cuando aplicamos un campo a una sustancia, ella desarrolla, para campos pequeños, un momento magnético proporcional al campo. Se llama *susceptibilidad* magnética al cociente entre M y B (para campos pequeños).

Ahora queremos considerar el paramagnetismo desde el punto de vista de la mecánica cuántica. Tomamos en primer lugar un átomo de espín $1/2$. En ausencia de campo magnético los átomos tienen una cierta energía, pero en un campo magnético hay dos energías posibles, una para cada valor de J_z . Para $J_z = +\hbar/2$, el campo magnético hace que la energía sufra una variación

$$\Delta U_1 = +g \left(\frac{q_e \hbar}{2m} \right) \cdot \frac{1}{2} \cdot B. \quad (35.10)$$

(El desplazamiento de energía ΔU es positivo para un átomo ya que la carga del electrón es negativa.) Para $J_z = -\hbar/2$, la energía varía en la cantidad

$$\Delta U_2 = -g \left(\frac{q_e \hbar}{2m} \right) \cdot \frac{1}{2} \cdot B. \quad (35.11)$$

Para ahorrar escritura, hagamos

$$\mu_0 = g \left(\frac{q_e \hbar}{2m} \right) \cdot \frac{1}{2}; \quad (35.12)$$

y entonces

$$\Delta U = \pm \mu_0 B. \quad (35.13)$$

El significado de μ_0 es claro: $-\mu_0$ es la componente z del momento magnético en caso de que el espín esté hacia arriba y $+\mu_0$ es la componente z del momento magnético en el caso de espín hacia abajo.

Ahora bien, la mecánica estadística nos dice que la probabilidad de que un átomo esté en un estado u otro es proporcional a

$$e^{-(\text{energía del estado})/kT}.$$

En ausencia del campo magnético, los dos estados tienen la misma energía y por eso cuando hay equilibrio en un campo magnético, las probabilidades son proporcionales a

$$e^{-\Delta U/kT}. \quad (35.14)$$

El número de átomos por unidad de volumen con espín hacia arriba es

$$N_{\text{arriba}} = ae^{-\mu_0 B/kT}, \quad (35.15)$$

y el número con espín hacia abajo es

$$N_{\text{abajo}} = ae^{+\mu_0 B/kT}. \quad (35.16)$$

La constante a hay que determinarla de manera que

$$N_{\text{arriba}} + N_{\text{abajo}} = N, \quad (35.17)$$

el número total de átomos por unidad de volumen. Así, obtenemos

$$a = \frac{N}{e^{+\mu_0 B/kT} + e^{-\mu_0 B/kT}}. \quad (35.18)$$

Lo que nos interesa es el momento magnético *promedio* según el eje z . Los átomos de espín hacia arriba contribuirán con un momento $-\mu_0$ y los de espín hacia abajo lo harán con $+\mu_0$; de este modo, el momento promedio es

$$\langle \mu \rangle_{\text{prom}} = \frac{N_{\text{arriba}}(-\mu_0) + N_{\text{abajo}}(+\mu_0)}{N}. \quad (35.19)$$

El momento magnético por unidad de volumen M es entonces $N\langle \mu \rangle_{\text{prom}}$. Usando las ecuaciones (35.15), (35.16) y (35.17) obtenemos

$$M = N\mu_0 \frac{e^{+\mu_0 B/kT} - e^{-\mu_0 B/kT}}{e^{+\mu_0 B/kT} + e^{-\mu_0 B/kT}}. \quad (35.20)$$

Ésta es la fórmula cuántica para M en el caso de átomos con $j = 1/2$. A propósito, esta fórmula también se puede escribir de un modo algo más conciso en términos de la función tangente hiperbólica:

$$M = N\mu_0 \tanh \frac{\mu_0 B}{kT}. \quad (35.21)$$

La figura 35-7 muestra una representación de M en función de B . Cuando B se hace muy grande, la tangente hiperbólica se aproxima a 1 y M hacia el valor límite $N\mu_0$. Por consiguiente, para campos intensos la magnetización se *satura*. Podemos ver por qué ocurre así; para campos suficientemente intensos todos los momentos están alineados en la misma dirección. En otras palabras, todos se encuentran en el estado de espín hacia abajo y cada átomo contribuye con el momento μ_0 .

En la mayoría de los casos normales —digamos, para los momentos típicos, la temperatura ambiente y los campos que comúnmente se puede obtener (del orden de 10 000 gauss)— el cociente $\mu_0 B/kT$ es del orden 0.002. Se debe bajar mucho la temperatura para ver la saturación. A temperaturas normales, se puede reemplazar $\tanh x$ por x y escribir

$$M = \frac{N\mu_0^2 B}{kT}. \quad (35.22)$$

Como en la teoría clásica, M es proporcional a B . De hecho, la fórmula es casi exactamente igual, excepto que parece que falta un factor de $1/3$. Pero aún necesitamos relacionar la μ_0 de la fórmula cuántica con la μ que aparece en el resultado clásico, ecuación (35.9).

En la fórmula clásica, lo que aparece es $\mu^2 = \boldsymbol{\mu} \cdot \boldsymbol{\mu}$, el cuadrado del vector momento magnético, o sea

$$\boldsymbol{\mu} \cdot \boldsymbol{\mu} = \left(g \frac{q_e}{2m} \right)^2 \mathbf{J} \cdot \mathbf{J}. \quad (35.23)$$

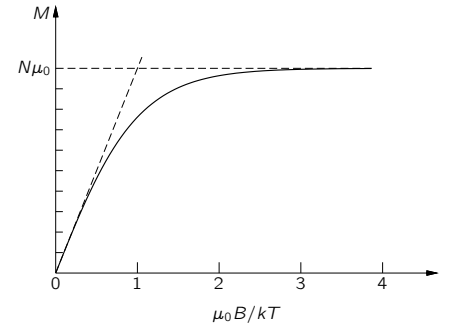


Figura 35-7. Variación de la magnetización paramagnética en función de la intensidad del campo magnético B .

Señalamos en el capítulo anterior que con mucha probabilidad pueden obtener la respuesta correcta de un cálculo clásico reemplazando $\mathbf{J} \cdot \mathbf{J}$ por $j(j+1)\hbar^2$. En nuestro caso particular tenemos que $j = 1/2$, por lo que

$$j(j+1)\hbar^2 = \frac{3}{4}\hbar^2.$$

Sustituyendo esto en lugar de $\mathbf{J} \cdot \mathbf{J}$ en la ecuación (35.23), obtenemos

$$\boldsymbol{\mu} \cdot \boldsymbol{\mu} = \left(g \frac{q_e}{2m}\right)^2 \frac{3\hbar^2}{4},$$

o en función de μ_0 , definida en la ecuación (35.12),

$$\boldsymbol{\mu} \cdot \boldsymbol{\mu} = 3\mu_0^2.$$

Sustituyendo μ^2 en la fórmula clásica (35.9) por esto, verdaderamente reproduce la fórmula cuántica correcta, ecuación (35.22).

La teoría cuántica del paramagnetismo se extiende fácilmente a átomos de cualquier espín j . La magnetización en un campo débil es

$$M = Ng^2 \frac{j(j+1)}{3} \frac{\mu_B^2 B}{kT}, \quad (35.24)$$

donde

$$\mu_B = \frac{q_e \hbar}{2m} \quad (35.25)$$

es una combinación de constantes con las dimensiones de un momento magnético. La mayoría de los átomos tienen momentos de este tamaño aproximadamente. Se llama *magnetón de Bohr*. El momento magnético de espín del electrón vale casi exactamente un magnetón de Bohr.

35-5 Enfriamiento por desmagnetización adiabática

Hay una aplicación especial del paramagnetismo que es muy interesante. A muy bajas temperaturas es posible alinear los imanes atómicos en un campo intenso. Es posible entonces descender a temperaturas *extremadamente* bajas mediante un proceso llamado *desmagnetización adiabática*. Podemos tomar una sal paramagnética (por ejemplo, una que contenga un cierto número de átomos de tierras raras como nitrato amónico de praseodimio), y comenzar a enfriarla con helio líquido hasta uno o dos grados absolutos en un campo magnético intenso. Así, el factor μ_B/kT es mayor que 1 —más parecido a 2 o 3—. La mayoría de los espines están alineados y la magnetización está casi saturada. Digamos, para simplificar, que el campo es muy potente y que la temperatura es muy baja, por lo que casi todos los átomos están alineados. Ahora aislamos la sal térmicamente (digamos que quitando helio líquido y haciendo un buen vacío) y apagamos el campo magnético. La temperatura de la sal sigue bajando.

Ahora bien, si quitasen el campo *de repente*, el movimiento desordenado y las sacudidas de los átomos en la red cristalina sacarían gradualmente todos los espines del alineamiento. Unos estarían hacia arriba y otros hacia abajo. Pero si no hay campo (y despreciando las interacciones entre los imanes atómicos, que serán solamente un error despreciable), no se necesita energía para voltear los imanes atómicos. Podrían distribuir al azar sus espines sin ningún cambio de energía y por lo tanto sin ningún cambio de temperatura.

Supongan, sin embargo, que mientras los imanes atómicos están siendo zaran-deados por el movimiento térmico, existe todavía un poco de campo magnético presente. Entonces se necesita trabajo para ponerlos en dirección opuesta al campo —*deben realizar trabajo contra el campo*—. Esto toma energía de los movimientos térmicos y baja la temperatura. Así pues, si el campo magnético intenso no se retira muy rápidamente, la temperatura de la sal disminuirá —se ha enfriado por desmagnetización—. Desde el punto de vista cuántico, cuando el campo es intenso todos los átomos están en el estado inferior, porque la

probabilidad de que alguno se encuentre en el estado superior es increíblemente pequeña. Pero a medida que disminuye el campo hay más y más probabilidad de que las fluctuaciones térmicas empujen un átomo al estado superior. Cuando esto sucede, el átomo absorbe la energía $\Delta U = \mu_0 B$. Por consiguiente, si el campo magnético se saca lentamente, las transiciones magnéticas pueden tomar energía de las vibraciones térmicas del cristal, enfriándolo. Así que es posible ir desde una temperatura de unos pocos grados absolutos hasta una temperatura de unas pocas milésimas de grado.

¿Les gustaría enfriar aún más? Pues bien, la naturaleza ha proporcionado un medio. Ya hemos mencionado que los núcleos atómicos también tienen momentos magnéticos. Nuestras fórmulas del paramagnetismo sirven igualmente bien para los núcleos, excepto que los momentos de los núcleos son aproximadamente un *millar de veces más pequeños*. [Son del orden de magnitud de $q\hbar/2m_p$ donde m_p es la masa del *protón*, por lo que son menores por el cociente de las masas del electrón y del protón.] Con tales momentos magnéticos, aun a una temperatura de 2°K, el factor $\mu B/kT$ es solamente unas pocas milésimas. Pero si usamos el proceso de desmagnetización paramagnética para descender a una temperatura de unas milésimas de grado, $\mu B/kT$ se hace casi 1 —a estas temperaturas bajas podemos comenzar a saturar los momentos nucleares—. Es buena suerte, ya que es posible usar entonces la desmagnetización adiabática del magnetismo *nuclear* para alcanzar temperaturas aún más bajas. En resumen, es posible realizar dos etapas para enfriamiento magnético. Primero usamos desmagnetización adiabática de iones paramagnéticos hasta alcanzar unas milésimas de grado. Luego, usamos la sal paramagnética fría para enfriar cierto material que tenga un magnetismo nuclear intenso. Finalmente, cuando quitamos el campo magnético de este material, su temperatura descenderá a unas *millonésimas* de grado sobre el cero absoluto —si lo hemos hecho todo muy cuidadosamente—.

35-6 Resonancia magnética nuclear

Hemos dicho que el paramagnetismo atómico es muy pequeño y que el magnetismo nuclear es aún un millar de veces menor. Incluso así es relativamente fácil observar el magnetismo nuclear mediante el fenómeno de “resonancia magnética nuclear”. Supongan que tomamos una sustancia, como agua, en la que cualesquiera de los espines electrónicos están exactamente equilibrados de manera que su momento magnético resultante es cero. Las moléculas aún tendrán un momento magnético muy, muy débil debido al momento magnético nuclear del núcleo de hidrógeno. Supongan que colocamos una muestra pequeña de agua en un campo magnético B . Como los protones (del hidrógeno) tienen espín 1/2, tendrán dos estados posibles de energía. Si el agua está en equilibrio térmico, habrá levemente más protones en los estados de energía más baja —con sus momentos paralelos al campo—. Hay un momento magnético pequeño resultante por unidad de volumen. Como el momento del protón es solamente una milésima de un momento atómico, la magnetización que va como μ^2 —usando la ecuación (35.22)— tiene solamente una millonésima de la intensidad del paramagnetismo atómico típico. (Ésta es la razón por la que hemos escogido un material sin magnetismo atómico.) Si lo calculan, la diferencia entre el número de protones con espín hacia arriba y con espín hacia abajo es solamente uno en 10^8 ; ¡el efecto es verdaderamente muy pequeño! No obstante, se puede observar de la siguiente manera.

Supongan que rodeamos la muestra de agua con una pequeña bobina que produce un pequeño campo magnético horizontal oscilante. Si el campo oscila a la frecuencia ω_p , inducirá transiciones entre los dos estados de energía —justamente como describimos en el caso del experimento de Rabi en la sección 35-3—. Cuando un protón salta de un estado de energía superior a uno inferior, suelta la energía $2\mu_z B$, que, como hemos visto, es igual a $\hbar\omega_p$. Si salta del estado de energía inferior al superior, *absorberá* la energía $\hbar\omega_p$ de la bobina. Como hay levemente más protones en el estado inferior que en el superior, habrá una *absorción* resultante

de energía de la bobina. Aunque el efecto es muy pequeño, la leve absorción de energía se puede ver con un amplificador electrónico sensible.

Lo mismo que en el experimento de haces moleculares de Rabi, la absorción de energía se verá solamente cuando el campo oscilante esté en resonancia, es decir, cuando

$$\omega = \omega_p = g \left(\frac{q_e}{2m_p} \right) B.$$

Normalmente es más conveniente buscar la resonancia variando B y conservando ω fijo. La absorción de energía aparecerá evidentemente cuando

$$B = \frac{2m_p}{gq_e} \omega.$$

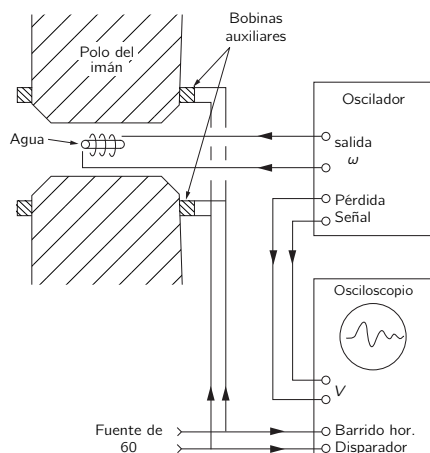


Figura 35-8. Un aparato de resonancia magnética nuclear.

La figura 35-8 muestra un típico aparato de resonancia magnética nuclear. Un oscilador de alta frecuencia se conecta a la bobina pequeña colocada entre los polos de un gran electroimán. Dos bobinas pequeñas auxiliares se conectan a una corriente de 60 ciclos de modo que el campo magnético se “tambalea” alrededor de su valor promedio en una cantidad muy pequeña. Como ejemplo, digamos que la corriente principal del imán se ajusta para producir un campo de 5 000 gauss y que las bobinas auxiliares producen una variación de ± 1 gauss respecto a ese valor. Si se coloca el oscilador a 21.2 megaciclos por segundo, se encontrará la resonancia protónica cada vez que el campo barra 5 000 gauss [usando la ecuación (34.13) con $g = 5.58$ para el protón].

El circuito del oscilador es de tal forma que da una señal de salida adicional proporcional a cualquier *variación* de potencia que se absorba del oscilador. Esta señal alimenta el amplificador de desviación vertical de un osciloscopio. El barrido horizontal del osciloscopio se dispara una sola vez en cada ciclo de la frecuencia tambaleante del campo. (Más comúnmente, la desviación horizontal se hace proporcional al campo tambaleante.)

Antes de que la muestra de agua se coloque en la bobina de alta frecuencia, la potencia sacada del oscilador tiene algún valor. (No varía con el campo magnético.) Sin embargo, cuando se coloca una botella de agua en la bobina, aparece una señal en el osciloscopio, como lo muestra la figura. Vemos una imagen de la potencia absorbida por la reorientación del espín de los protones.

En la práctica, es difícil saber cómo ajustar el imán principal a exactamente 5 000 gauss. Lo que se hace es ajustar la corriente del imán principal hasta que la señal de resonancia aparece en el osciloscopio. Resulta que ésta es ahora la forma más conveniente de hacer una medida precisa de la intensidad de un campo magnético. Desde luego, antes *alguien* tuvo que medir muchas veces y con precisión el campo magnético y la frecuencia para determinar el valor de g del protón. Pero una vez que se hace esto, un aparato de resonancia protónica como el de la figura se puede usar como “magnetómetro de resonancia protónica”.

Diremos algo de la forma de la señal. Si variásemos el campo magnético muy lentamente, esperaríamos ver una curva de resonancia normal. La absorción de energía sería máxima cuando ω_p igualase exactamente la frecuencia del oscilador. Habría alguna absorción a frecuencias cercanas ya que todos los protones no se encuentran en el mismo campo exactamente —y campos diferentes significan frecuencias de resonancia levemente distintas—.

Incidentalmente, uno se podría preguntar si a la frecuencia de resonancia deberíamos ver alguna señal. ¿No sería de esperar que el campo de alta frecuencia igualase la población de los dos estados —así no habría señal excepto en el momento de colocar el agua—? No exactamente, porque aun cuando *tratemos* de igualar las dos poblaciones, los movimientos térmicos por su parte tratan de conservar la proporción para la temperatura T . Si nos situamos en la resonancia, la potencia que los núcleos absorben es justamente la que se pierde por los movimientos térmicos. Sin embargo, hay relativamente poco “contacto térmico” entre los momentos magnéticos protónicos y los movimientos atómicos. Los protones están relativamente aislados en el centro de las distribuciones electrónicas. Por lo tanto, la señal de resonancia en agua pura es en realidad usualmente muy

pequeña para ser vista. Para aumentar la absorción es necesario incrementar el “contacto térmico”. Esto se hace comúnmente añadiendo al agua un poco de óxido de hierro. Los átomos de hierro son como imanes pequeños; cuando se zarandean en su danza térmica, producen campos magnéticos danzantes pequeñísimos en los protones. Estos campos variables “acoplan” los imanes protónicos a las vibraciones atómicas y tienden a establecer el equilibrio térmico. Es por este “acoplamiento” que los protones en los estados de energía más alta pueden perder su energía y así son capaces de absorber nuevamente energía del oscilador.

En la práctica la señal de salida de un aparato de resonancia nuclear no se parece a una curva de resonancia normal. Habitualmente es una señal más complicada con oscilaciones —como la dibujada en la figura—. Tales señales aparecen debido a los campos variables. La explicación de ello debería ser cuántica, pero se puede demostrar que en esos experimentos las ideas clásicas de momentos que precesan dan siempre la respuesta correcta. Clásicamente diríamos que cuando llegamos a la resonancia comenzamos a tener un montón de imanes nucleares que precesan en sincronía. Al realizarlo, hacemos que precesen *juntos*. Estos imanes nucleares, todos rotando juntos, inducirán una fem en la bobina del oscilador a la frecuencia ω_p . Pero como el campo magnético aumenta con el tiempo, la frecuencia de precesión también aumenta y el voltaje inducido se encuentra pronto a una frecuencia un poco mayor que la frecuencia del oscilador. Cuando la fem inducida esté alternadamente en fase o desfasada con el oscilador, la potencia “absorbida” será positiva o negativa alternadamente. Así, en el osciloscopio vemos pulsación entre la frecuencia protónica y la frecuencia del oscilador. Como las frecuencias de los protones no son todas idénticas (protones diferentes se encuentran en campos ligeramente distintos) y también posiblemente debido a la perturbación del óxido de hierro en el agua, los momentos que precesan libremente pronto se desfasan y la pulsación desaparece.

Estos fenómenos de resonancia magnética se han usado de diversos modos como herramientas para nuevos hallazgos en la materia —especialmente en química y en física nuclear—. Está demás decir que los valores numéricos de los momentos magnéticos de los núcleos nos indican algo de su estructura. Se ha aprendido mucho en química con la estructura (o forma) de las resonancias. A causa de los campos magnéticos producidos por núcleos cercanos, la posición exacta de la resonancia nuclear está algo movida dependiendo del ambiente en que un núcleo particular se encuentra. La medida de este desplazamiento ayuda a determinar cuáles átomos están cerca de cuáles y ayuda a esclarecer los detalles de la estructura molecular. Igualmente importante es la resonancia de espín electrónico de radicales libres. Aunque no haya muchos de esos radicales en el equilibrio, sí son con frecuencia estados intermedios de reacciones químicas. Una medida de una resonancia de espín electrónico es una comprobación delicada de la presencia de radicales libres y es a menudo la clave para entender el mecanismo de ciertas reacciones químicas.

Índice analítico

A

Absorción
 de fotones, 3, 4-1, 4-8
 de luz, 3, 9-1, 9-15
Adjunta, 11-24
Adjunto hermitiano, 20-3
Álgebra de matrices, 5-15, 11-3, 20-18
Álgebra vectorial, 5-16, 8-2, 8-4
Amplitudes, 3, 8-1
 Dependencia espacial de las, 4, 13-4, 16-1
 Dependencia temporal de las, 3, 7-1
 Transformación de, 6-1
Amplitudes de probabilidad, 1-10, 3-2, 16-1
Amplitudes interferentes, 5-10
Amplitudes transformadoras, 3, 6-1
Antimateria, 11-16
Antipartícula, 11-14
Antiprotón, 11-14
Aparato de Stern-Gerlach, 5-12
Aproximación de partículas independientes, 4, 15-1
Argón, 19-17
Armónicos esféricos, 19-8, 19-9
Átomo de hidrógeno, 4, 19-1

B

Banda de conducción, 14-1, 14-4
Barion, 11-14
Boro, 19-16
Bosones, 4-2, 15-7

C

Celda unitaria, 2-5
Coeficiente
 de Clebsch-Gordan, 18-16, 18-17, 18-20
 de Einstein, 9-16
Colorante, 10-12
Configuración electrónica, 19-15
Conservación
 de extrañeza, 11-13
 de la energía, 3, 7-1, 7-6
 de número de bariones, 11-14

Constante
 de Boltzmann, 14-4
Constante de Planck, 2-3, 2-6, 20-15, 21-2
Criptón, 19-16
Cuantificación del flujo, 5, 21-1, 21-10

D

Densidad
 de carga eléctrica, 21-2
 de probabilidad, 16-8
Densidad de probabilidad, 16-6
Diagrama de energía, 14-2
Diagrama del nivel de energía, 14-3
Difracción
 de rayos X, 2-5
Difracción de cristales, 3, 2-1, 2-5
Distribución
 Normal (Gaussiana), 16-8
Distribución de probabilidad, 16-8
Distribución normal, 16-8
Divisiones hiperfinas en hidrógeno, 12-1

E

Ecuación, 4, 19-1
 de Schrödinger, 4, 16-1, 16-4, 16-12, 16-14, 21-1
 en el contexto clásico, 5, 21-1
Efecto
 Doppler, 2-6, 12-9
 Hall, 4, 14-1, 14-7
 Meissner, 5, 21-1, 21-8, 21-10
 Zeeman, 12-12
Electrón, 1-1, 1-7
Electrón sigma, 12-3
Elemento de matriz dipolar eléctrica, 9-16
Emisión de fotones, 3, 4-1, 4-8
Energía
 potencial, 3, 7-1, 7-6
Energía de ionización
 energía de ionización del hidrógeno, 2-7
Energía de Rydberg, 10-4, 19-3
Espectro de cuerpo negro, 3, 4-1, 4-9
Estado base, 5-8, 12-1
 del mundo, 3, 8-1, 8-5

Estado fundamental, 7-2
Estados
 de energía definida, 4, 13-1, 13-3
 Estacionarios, 11-18
 excitados, 13-9
 propio, 11-23
 Tiempo dependiente, 4, 13-1, 13-7
Excitón, 13-10
Extrañeza, 11-13
 conservación de, 11-13

F

Factor de Boltzmann, 14-4
Fotón, 1-8
 estados polarizados, 11-13
Fuerza
 centrífuga, 19-12, 19-14
 nuclear, 3, 10-1, 10-6
Función de onda, 4, 16-1, 16-5
 significado de la, 5, 21-1, 21-6
Funciones de Legendre, asociadas, 19-9
Funciones de onda de hidrógeno, 19-8

G

Galio, 19-18

H

Hamiltoniano, 10-2
Helio, 19-14
Helio líquido, 4-12
Hidrógeno, 4, 12-1, 19-1
 estructura hiperfina en el, 4, 12-1
Hidrógeno molecular, 10-2

I

Interacción espín-órbita, 15-14
Interferencia de dos rejillas, 3, 3-1, 3-5

J

Juntura de Josephson, 5, 21-1, 21-14
Junturas de semiconductores, 4, 14-1, 14-8
Rectificación en una, 4, 14-1, 14-11

L

Láser, 9-14
Ley de Ohm, 14-7
Límite clásico, 3, 7-1, 7-9
Línea de veintiún centímetro, 12-9
Litio, 19-15

M

Magenta, 10-13
magnética, 5-1
Magnetón de Bohr, 12-12
Magnons, 15-4
Masa
 efectiva, 13-7
Máser de amoníaco, 9-1
Matrices de espín de Pauli, 4, 11-1
Matrices sigma, 11-2
Matriz, 5-5
Matriz de rotación, 6-4
Matriz hamiltoniana, 3, 8-1, 8-10
Matriz unitaria, 11-2
Mecánica cuántica, 1-1, 2-1, 3-1,
 5-1, 21-2
 del potencial vectorial, 21-2
Mesón K neutro, 4, 11-1, 11-13
Molécula de amoníaco, 8-14
 estados de una, 3, 9-1
Molécula de benceno, 10-11, 15-8
Molécula de clorofila, 15-12
Molécula de etileno, 15-8
Molécula de hidrógeno ionizada, 10-2
Molécula de trifenil ciclopropenilo,
 15-13
Momentos magnéticos, 11-4
Momentum, 3, 2-1, 2-2
 angular, 4, 18-1, 20-1, 20-14
 Composición de, 4, 18-1, 18-15
Momentum angular orbital, 19-9
Momentum cinemático (mv), 21-5
Momentum dinámico (p), 21-5

N

Neón, 19-15
Niveles de energía, 3, 2-1, 2-7
Nodos de onda, 7-10
Nucleón, 11-3
Número cuántico principal, 19-13
Números cuánticos, 12-15

O

Ondas, 4, 15-1
 de espín, 4, 15-1
Ondas interferentes, 1-4
Operador, 4, 20-1
 algebraico, 20-3
 de intercambio de espín de Pauli,
 12-7, 15-2
 momentum, 4, 20-1, 20-3, 20-9

P

Paquete de onda, 13-7
Paradoja de Einstein-Podolsky-Rosen,
 18-9
Partículas, 3, 4-1, 7-1, 7-11
 bosones, 15-7
 espín un medio, 3, 6-9, 7-1,
 7-11, 12-3
 La precesión de, 3, 7-1, 7-11
 espín uno, 5-2
 fermión, 4-2, 15-7
 Idénticas, 4-1
Pion neutro, 10-7
Polinomios de Legendre, 18-13, 19-9
Portadores, 14-2
 Negativo, 14-2
 Positivo, 14-2
Portadores negativos, 14-4
Portadores positivos, 14-4
Potasio, 19-17
Potencial
 de Yukawa, 10-7
 vectorial
 y la mecánica cuántica, 21-2
Principio, 4-14
 Exclusión, 4-14
Principio de combinación de Ritz, 2-8
Principio de exclusión, 4-14
Principio de incertidumbre, 1-9, 1-11,
 1-12, 2-6, 2-9
Procesos atómicos
 Consejación de la paridad, 18-2
Propagación en una red cristalina,
 4, 13-1

R

Radio, 2-7
 de Bohr, 2-7, 19-3, 19-6

Receptor, 14-5
Red cristalina
 imperfección del, 13-10
 Propagación en, 13-5
Red de diamante, 14-1
Red tridimensional, 4, 13-1, 13-8
Red unidimensional, 4, 13-1
Regla de conmutación, 20-15
Resonancia, 5-1, 10-4
 cuántica, 10-4
Resonancia magnética cuántica, 10-16
Rydberg (unidad), 2-7

S

Semiconductores, 4, 14-1
 Impuros, 4, 14-1, 14-4
 tipo n , 14-5
 tipo p , 14-6
Separación de Zeeman, 4, 12-1, 12-9
Sigma protón, 12-4
Sistemas de dos estados, 10-1, 11-1
Sitio donador, 14-5
Sodio, 19-17
Soluciones con simetría esférica, 4,
 19-1, 19-2
Superconductividad, 5, 21-1

T

Tabla periódica, 4, 19-1, 19-14
Tensor algebraico, 8-4
Teorema
 de Gauss, 21-4
Teoría cinética
 Aplicaciones de, 4-1
Teoría de Boltzmann, 21-8
Transistores, 14-12

V

Valores propios, 11-23
Vector de estado, 8-1
 Descomposición de, 3, 8-1, 8-3
Vector sigma, 11-4
Visión, 13-10

Z

Zinc, 19-18

Índice de nombres

B

Bohr, Niels (1885–1962), 16-14,
19-3, 19-6
Born, Max (1882–1970), 1-1, 2-9,
3-1, 21-6

D

Dirac, Paul A. M. (1902–1984), 3-1, 8-2,
12-7, 16-10, 16-11, 16-14

E

Einstein, Albert (1879–1955), 4-9, 18-9

G

Gell-Mann, Murray (1929–2019), 11-13,
11-17–11-20

Goeppert-Mayer, Maria (1906–1972),
15-14

H

Hamilton, William Rowan (1805–65),
8-11
Heisenberg, Werner K. (1901–76), 1-1,
1-11, 1-12, 2-9, 16-9,
20-17, 20-18

J

Jensen, J. Hans D. (1907–1973), 15-14
Josephson, Brian D. (1940–), 21-14,
21-16–21-18

N

Newton, Isaac (1643–1727), 1-1

Nishijima, Kazuhiko (1926–2009), 11-13

P

Pais, Abraham (Bram) (1918–2000),
11-13, 11-17–11-20
Pauli, Wolfgang E. (1900–1958), 4, 4-3,
11-1, 11-2, 11-5
Planck, Max (1858–1947), 4-12

S

Schrödinger, Erwin (1887–1961), 4, 5,
1-1, 16-1, 16-4, 16-12–16-15,
20-8, 20-17, 20-18, 21-1–21-8,
21-12, 21-14

Y

Yukawa, Hideki (1907–1981), 10-6, 10-7

Lista de símbolos

$\langle f s \rangle$	amplitud para que un sistema preparado en el estado inicial $ s\rangle$ se encuentre en el estado final $ f\rangle$. 3-2
$Y_{l,m}(\theta, \phi)$	armónicos esféricos. 19-8
C	coeficientes de Clebsh-Gordan. 18-20
κ	constante de Boltzmann, $\kappa = 1.3806504 \times 10^{-23}$ J/k. 14-4
E_{dif}	diferencia de energía. 14-4
$\mathbf{n}$	el n -ésimo número Romano, de modo que $\mathbf{n}$ toma los valores $I, II, \dots, \mathbf{N}$. 11-22
$ L\rangle$	estado de fotones polarizados circularmente a la izquierda. 11-11
$ R\rangle$	estado de fotones polarizados circularmente a la derecha. 11-11
$\langle \phi $	estado ϕ escrito como un vector “bra”. 8-2
$ \phi\rangle$	estado ϕ escrito como un vector “ket”. 8-2
$\mathcal{I}$	intensidad. 9-15
m_{ef}	masa efectiva de electrones en una red cristalina. 13-7
$\sigma_x, \sigma_y, \sigma_z$	matrices giratorias de Pauli. 11-2
N_p	número de agujeros por unidad de volumen. 14-4
N_n	número de electrones por unidad de volumen. 14-4
$P_{\text{int espín}}$	operador de cambio de giro de Pauli. 12-7
$U(t_2, t_1)$	operador que designa la operación esperando desde el tiempo t_1 hasta t_2 . 8-8
$ 1\rangle, 2\rangle$	una elección específica de base de vectores para un sistema de dos estados. 9-1
$ I\rangle, II\rangle$	una elección específica de base de vectores para un sistema de dos estados. 9-3
Φ_0	unidad básica de flujo. 21-12
$\mathcal{E}$	vector de campo eléctrico. 9-5
$\mathcal{E}_{\text{tr}}$	vector de campo eléctrico transversal. 14-8
σ	vector de matrices giratorias de Pauli. 11-4

Los textos contemporáneos de física pueden catalogarse en un antes y un después de las lecciones de Richard Feynman. Propuestas a partir de las clases impartidas en el Instituto Tecnológico de California (Caltech) de 1961 a 1963, Feynman transformó la concepción de esta pequeña ventana al universo con una nueva propuesta llena de vitalidad en su enfoque. A partir de entonces, la enseñanza de esta ciencia logra profundizar el estudio de los fenómenos físicos de forma puntual y sencilla.

Feynman, Premio Nobel de Física en 1965, brinda en este tercer y último volumen un primer acercamiento a los fundamentos de la mecánica cuántica; es decir, del comportamiento de la materia y la energía a escala atómica. Además de explorar sus principios básicos, revisa la estructura matemática de las teorías cuánticas, útil para resolver problemas que se presentan en otros campos de la física. Este volumen deja en claro cuánto se puede deducir de las dimensiones a escala atómica.

La física de Feynman es sobre simplicidad, belleza, unidad y analogía, presentada con entusiasmo y perspicacia.

Nature

Una reacción en cadena no es una mala analogía para la vida de Feynman. A partir de una masa crítica de materia gris se dispara en todas direcciones, produciendo tanto calor como luz.

Time