ANEXOS

ANEXO 1

EL EQUILIBRIO EN EL ANALISIS DEL INGRESO NACIONAL

Aunque en el estudio del análisis estático, visto en el Capitulo II, el modelo de mercado de varios tipos (lineales y no lineales, de un articulo o de muchos, específicos y generales) tiene aplicación también en otros campos de la economía. Vrg. el conocido modelo Keynesiano del Ingreso Nacional.

$$Y = C + Io + Go \qquad (1)$$

$$C = a + bY \qquad (2)$$

Donde Y y C representan las variables endógenas del ingreso nacional y gastos de consumo, respectivamente, e Io y Go representan las inversiones y gastos en las que incurre el gobierno, determinados exógenamente. La primera ecuación es una condición de equilibrio (Ingreso Nacional = Gasto Total). La segunda, que es la función consumo, es de comportamiento.

Esta muy claro que estas dos ecuaciones con dos variables endógenas o son funcionalmente dependientes ni recíprocamente incompatibles. Así podemos hallar los valores de equilibrio de ingreso y gastos de consumo, Y' y C', en términos de los datos dados:

"a" y "b" son parámetros dados.

Io y Go son las variables exógenas.

Al sustituir la segunda ecuación en la primera, reduciremos a una única ecuación en una variable, Y:

$$Y = a + bY + Io + Go$$

$$(1 - b)Y = a + Io + Go$$

Y así, el valor de solución de Y (Ingreso Nacional de Equilibrio) es:

$$Y' = \frac{a + Io + Go}{1 - b} \qquad (*)$$

Se advierte que esto se expresa enteramente en términos de los parámetros y variables exógenas que esta en los datos del modelo.

Si ponemos (*) en la ecuación (2), entonces obtendremos el nivel de equilibrio de gastos de consumo, de la siguiente manera:

$$C' = a + bY' = a + b\left(\frac{a + Io + Go}{1 - b}\right) = \frac{a + b(Io + Go)}{1 - b} \qquad (**)$$

Tanto Y' como C' tienen la expresión (1-b) en el denominador; lo que hace necesaria la restricción $b \neq 1$. Además, b representa la Propensión Marginal a Consumir, que por lo general es una fracción positiva $(0 \prec b \prec 1)$, esta restricción queda automáticamente satisfecha. Para que Y' y C' sean positivas, además, los numeradores de (*) y (**) deben ser positivos.

Puesto que los gastos exógenas Io y Go son normalmente positivos, como lo es el parámetro "a" (la intersección vertical de la función consumo –se dirá, consumo autónomo-), también corresponderá el signo de las expresiones del numerador.

Como prueba de nuestros cálculos podemos sumar la expresión C' en "*a(Io+Go)*" y comprobar si la suma es igual a la expresión Y' en (*). Si así fuere, C' y Y' satisfarían la condición de equilibrio y la solución es valida, como se quería demostrar.

ANEXO 2

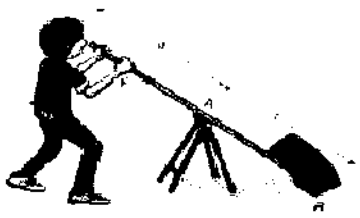
"Curso De Física" Traducción de la 8va. Ed. Inglesa. Editorial LABOR 1939. W. Watson,

1. PALANCAS:

La palanca es una barra rígida que puede girar alrededor de un punto fijo o eje, cuando se le aplica una fuerza para vencer la resistencia. Es usada para vencer la mayor resistencia con el esfuerzo aplicado. Las palancas se clasifican en 1ª,2ª y 3ª clase, esto se hace según la ubicación del punto de apoyo A(llamado fulcro), de la fuerza motora F y de la resistencia. Haremos también una distinción de la "palanca matemática" (en la que no se considera el peso de la palanca) y la "palanca física" (en la que sí se considera el peso de la palanca).

Esta máquina es una aplicación directa del principio de momentos donde una de las fuerzas (la resistencia R) hace girar la palanca en un sentido. Es decir, constituye un momento $R \times r$ respecto de A; mientras que la otra fuerza (motriz F), en el sentido contrario, tiene un momento de inercia determinado por $F \times d$ (d y r son los brazos respectivos).

La condición para que una palanca se mantenga en equilibrio es que la suma de los momentos de fuerza motriz y de la resistencia sean nulas.



Esto es: $\Sigma \tau_i = 0$

O sea: $F \times d = R \times r$ lo cual representa la ley de la palanca.

d = Su valor debe estar en Metros
r = Su valor debe estar en Metros

"El producto de la fuerza por su brazo es igual al producto de la resistencia por su brazo"

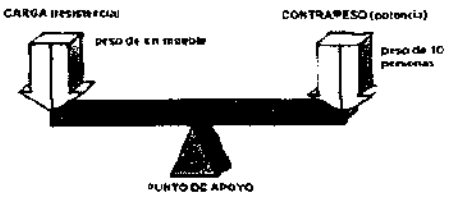
Palancas de Primera clase:

Como dijimos al principio, son aquellas en la cuales su punto de apoyo está entre la resistencia y la fuerza motriz.

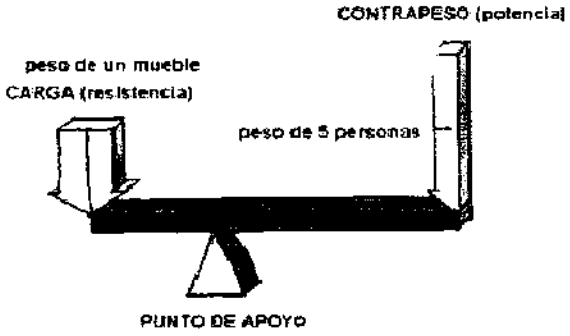
El equilibrio se da porque la fuerza F da un momento positivo **M+** $= F \times d$ que se anula con el momento negativo de la resistencia R **M-** $= R \times r$. Al producirse esta estabilización de momentos, la ecuación de equilibrio es la mencionada anteriormente. ($F \times d - R \times r$)

De esta ecuación podemos calcular el valor de F necesario para poder equilibrar una, a través de la siguiente: $F = \frac{R \times r}{d}$

Con la ecuación mostrada anteriormente para obtener el valor de F se puede concluir que para obtener una mayor economía de fuerza conviene que el brazo de la fuerza motora sea grande y el de la resistencia pequeño.



Cuando el punto de apoyo está en la mitad, la longitud entre el punto medio y F, es igual a la longitud entre el punto medio y R, por lo que para que se mantenga en equilibrio las fuerzas deben ser iguales.



Cuando el punto de apoyo está descentrado, se da el caso en que se necesita una fuerza menor para poder equilibrar la balanza, con esto la ventaja mecánica es ahorro de fuerza (contrapeso). En la siguiente figura se muestra que la distancia entre el apoyo y F es el doble que la distancia entre el apoyo y R, por lo que d = 2r, con esto si lo llevamos a la ecuación obtenemos F = 2R , es decir necesitamos la mitad de la fuerza que ejerce la carga para mantener en equilibrio la palanca.

Ejemplo:

- Se tiene una palanca de 4m de largo en la que hay una carga de 20kg, que está a 2,7m del eje. ¿Cuál será el valor de la fuerza motora si esta fuerza se encuentra a 1.3 m del eje?. El peso de la barra es despreciable.

Datos :

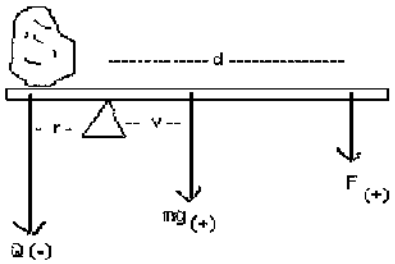
$F = ¿?$
$d = 1.3m$
$R = 200N$
$r = 2.7m$

Aplicamos la fórmula dicha anteriormente:

$$F \times d = R \times r \rightarrow F = \frac{200N \times 2.7m}{1.3m} = 415.3846154N$$

Ejemplo de palanca física de 1ª clase. (Considerando el peso de la palanca).

Consideraremos la barra como un material grueso y uniforme, es decir que su centro de gravedad 'G' está en el centro de ella donde actúa su peso **P= mg** con un brazo "v".



Habíamos quedado en que la Resistencia producía un momento negativo igual a $R \times r$ que era equilibrado por el momento positivo de $F \times d$, pero ahora le agregaremos el peso de la palanca (P = mg), con lo queda la siguiente expresión: $F \times d + mg \times v = R \times r$

En la figura se puede ver que el peso está a favor de la fuerza

Ejemplo:

- Para mover una piedra (Q) que hace una fuerza de 20kg(en la barra hace una fuerza resistente de 200N) se usa una barra de 1.6 m de largo y de 1 kg (osea hace una fuerza de 10N) de peso, como se muestra en la figura. ¿Dónde debe situarse el punto de apoyo si en el extremo de la barra una persona hace una fuerza de 6kg.?

Desarrollo:

r = x = el brazo de la resistencia
v = 80 – x = el brazo del peso (ya que G es el punto medio de la barra)
d = 160 – x = el brazo de la fuerza
Ahora realizamos la ecuación igualando los momentos positivos con los negativos y queda:
60 * (160 - x) + 10 * (80 – x) = 200x
Resolviendo la ecuación nos queda x = 38.5 cm → 0.385 cm

Para este ejercicio y para algunos de más adelante, se considerará lo siguiente: Si un cuerpo pesa 10Kg, será tomado como una fuerza de 100N, ya que le multiplicaremos la aceleración que experimenta (a = 10m/s²)

Ahora mostraremos unos ejemplos de palancas de primera clase.

carretilla para poder transportar una pesada carga con un pequeño esfuerzo.

Palancas de Segunda clase:

En este tipo de palancas el punto de apoyo está en un extremo de ella, la potencia o fuerza motriz (F) en el otro extremo y la resistencia (R) en algún punto intermedio. Si vemos el dibujo, podemos llegar a las siguientes conclusiones:

M+ $-F \times d$, la resistencia R da un momento positivo.
M- $-R \times r$, la fuerza motriz produce un momento negativo.

Luego para que se quede en equilibrio tenemos que $F \times d = R \times r$.

Podemos ver como la fórmula obtenida es exactamente la misma que para una palanca de primer género. Así si tenemos un brazo suficientemente grande se puede vencer una resistencia grande aplicando una potencia pequeña.

Ejemplo:

- Suponga una carretilla de masa despreciable que mide 1m, que está cargando un libro de física que pesa 2 kg. Determine la distancia que hay entre el punto de apoyo y la carga, si la fuerza aplicada sobre el estudiante es de $10N$..

Datos:

$F = 10N$
$d = 1m$
$R = 20N$
$r = ¿?$

De la ecuación: $F \times d = R \times r$ obtenemos: $r = \frac{F \times d}{R}$ →

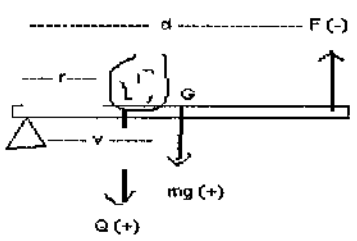
r = ½ → r = 0.5m

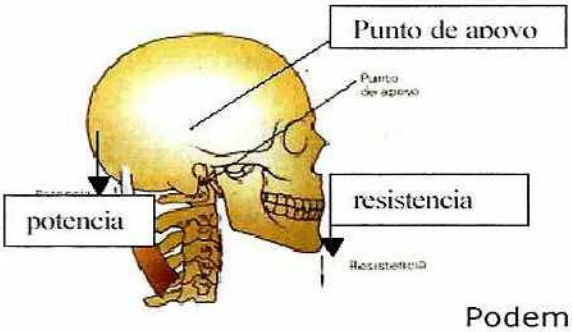
Ejemplo de palanca física de 2ª Clase.

Anteriormente habíamos quedado que la Resistencia producía un momento positivo (que ahora se le suma el peso de la barra), que era equilibrado por el momento negativo de la fuerza motora.

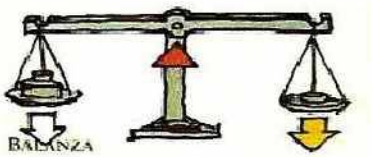
Entonces quedamos con la siguiente ecuación:

$$F * d = R * r + mg * v .$$

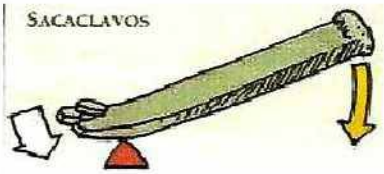




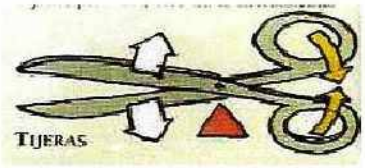
Podemos ver que el cuerpo humano está lleno de palancas de distintas clases, por ejemplo, la cabeza, donde su peso es contrarrestado por la acción de la musculatura de la nuca, tomando la columna vertebral como punto de apoyo.



También la 'típica' balanza que se encarga de pesar la carga, ésta debe reunir por lo menos tres condiciones que son:
Ser sensible (que cuando en uno de sus platillos le coloque un peso debe perder su equilibrio, por muy pequeña que sea la carga), también debe ser exacta (que al colocar pesos iguales en los platillos debe recuperar su posición de equilibrio) y finalmente debe ser estable (debe volver a la posición de equilibrio)



Este instrumento, suele estar en los martillos, aquí la fuerza la realiza la persona que intenta sacar el clavo. La carga es la resistencia del clavo al ser extraído.



Un material de uso diario, las tijeras. Son dos palancas combinadas de primer grado. Estas realizan un fuerte acción de corte cerca del punto de apoyo. La carga es la resistencia del material a la acción de corte de las hojas de las tijeras.

Finalmente "la carretilla de dos ruedas", usada para trasladar los balones de gas, entre otros. Basta inclinar las varas de la

Como se muestra en la imagen el peso ahora juega en contra de la fuerza

- Se tiene una piedra de 20kg(Q), que se encuentra a 20cm del punto de apoyo, si se usa una barra de 1m de largo, de 1kg de peso, como palanca de 2º grado, ¿Cuál debe ser el valor de la fuerza para equilibrar la balanza?

Datos:

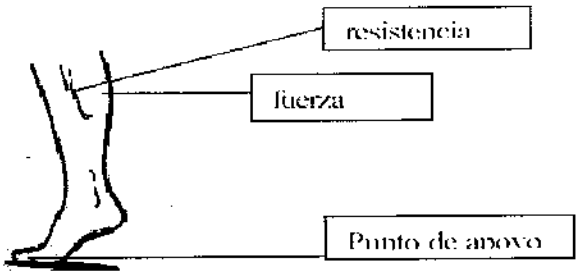
F = ¿?
d = 1m
R = 200N
mg = 10N
v = 0.5m.
r = 0.2m

Desarrollo:

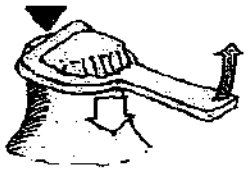
$F * d = R * r + mg * v \rightarrow F = \frac{(R * r + mg * v)}{d}$

$\rightarrow F = \frac{(40 + 5)}{1} \rightarrow F = 45\ N.$

Ejemplo de palancas de 2ª clase

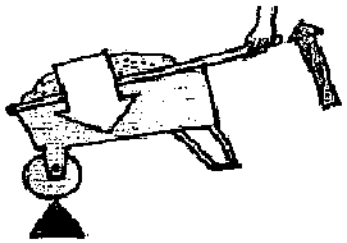


Cuando caminamos, nuestro pie se apoya en los dedos al caminar: La fuerza motora la hace el músculo de la pantorrilla y la resistencia es el peso del cuerpo.



ABRIDOR DE BOTELLAS

Últimamente hemos visto la utilidad semanal de este instrumento, el "abridor de botellas", al levantar el mango, se supera la fuerte resistencia de la tapa.



CARRETILLA DE UNA RUEDA

Finalmente, la carretilla, Al levantarla de las varas es posible levantar una pesada carga que se halla más cerca del punto de apoyo, la rueda.

ANEXO 3

Notas de clase sobre Equilibrio General. Nota 2: Repaso de Teoría del Consumidor

Andrés Carvajal
Universidad del Rosario, Departamento de Economía. acarvaes@banrep.gov.co

September 4, 2003

Abstract

Se presenta aquí una revisión breve de teoría del consumidor. Esta revisión está lejos de ser comprensiva. De hecho, sólo se abordan aquí aspectos que van a sernos útiles posteriormente en el curso.

Un consumidor es un caso particular de un agente que toma decisiones. En general, un tomador de decisiones se modela como tres elementos $(X, \succsim, \mathcal{B})$, donde X representa el espacio en el que el agente puede tomar decisiones, $\succsim$ es una relación binaria en X, según la cual el agente toma sus decisiones y $\mathcal{B}$ es una familia de subconjuntos de X, que es la familia de problemas que el agente enfrenta.

Más explícitamente:

- $X \neq \varnothing$ es el espacio de elección del agente: el conjunto de todas las alternativas que el agente podría, concebiblemente, elegir, más allá de retricciones de factibilidad o de sus gustos.
- $\succsim$ es una relación binaria en X (es decir un subconjunto de $X \times X$): si $x, x' \in X$, $x \succsim x'$ quiere decir que el agente encuentra a x al menos tan bueno como x'.
- Si $B \in \mathcal{B}$, quiere decir que el agente puede enfrentar el problema de escoger $x \in B \subseteq X$.

El supuesto de comportamiento es que cuando el agente tiene que escoger en $B \in \mathcal{B}$, escoge un elemento que es máximo con respecto a $\succsim$: el agente escoge $x \in B$ si no existe $x' \in B$ tal que $x' \succsim x$ sin que $x \succsim x'$.[1] Para que este problema esté bien definido, es necesario hacer algunos supuestos sobre estos tres objetos. Dado que nuestro objetivo es estudiar consumidores, loq ue baremos a continuación es hacer estos supuestos para el caso concreto de problemas de elección del consumidor.

Vamos a suponer aquí que existen $L \geqslant 1$ bienes. Por bien entendemos cualquier objeto, tangible o intangible que es sujeto de intercambio. Implícitamente, estamos suponiendo que nuestros agentes conocen los bienes con mucha precisión y que son capaces de medirlos. Así, definimos $X = \mathbb{R}^L_+$, es decir, suponemos que el espacio de elección es el espacio de canastas de los L bienes, consumidos en cantidades no negativas. Suponemos además que:

Supuesto 1 *$\succsim$ es racional:*

1. *Es completa:* $\forall x, x' \in \mathbb{R}^L_+$, $x \succsim x'$ *ó* $x' \succsim x$.
2. *Es transitiva:* $\forall x, x', x''$, *si* $x \succsim x'$ *y* $x' \succsim x''$, *entonces* $x \succsim x''$.

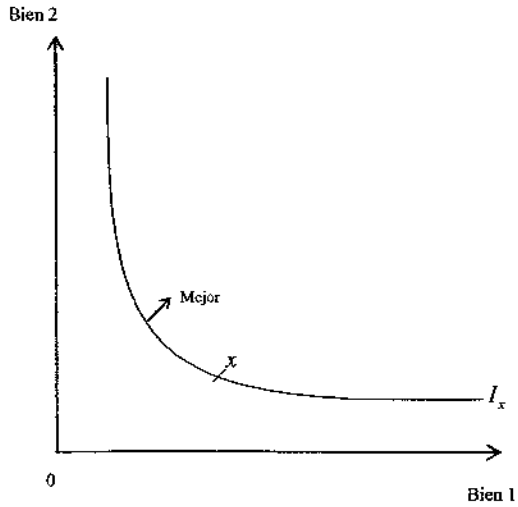
[1] La siguiente notación será utilizada: $x \sim x'$ querrá decir que $x \succsim x'$ y $x' \succsim x$; $x \succ x'$ querrá decir que $x \succsim x'$ pero que no es cierto que $x' \succsim x$.

Esto hace que $\succsim$ sea un preorden total. Completitud es relativamente razonable: si los agentes conocen bien los bienes como hemos supuesto, entonces ellos deberían saber qué les gusta más y qué menos. El supuesto de transitivaidad es más controversial pues hay evidencia experimental de que los seres humanos habitualmente lo violamos, pero resulta fundamental para que la teoría tenga piso: a un consumidor no transitivo se le pueden extraer rentas infinitas.

Ejercicio 1 *Demuestre lo siguiente: si $\succsim$ es racional entonces:*

1. $\succ$ *es irreflexiva:* $\forall x \in \mathbb{R}^L_+$, *no es cierto que* $x \succ x$.
2. $\succ$ *es transitiva.*
3. $\sim$ *es reflexiva:* $\forall x \in \mathbb{R}^L_+$, *es cierto que* $x \sim x$.
4. $\sim$ *es transitiva.*
5. $\sim$ *es simétrica: si* $x \sim x'$, *entonces* $x' \sim x$.

Ahora, en economía estamos acostumbrados a trabajar sobre la base de curvas de indiferencia en el espacio de consumos que tienen una forma muy habitual:



En el lenguaje que estamos utilizando ahora, la curva de indiferencia que hemos dibujado es

$$\{x' \in X | x' \sim x\}$$

Entonces, la forma que habitualmente suponemos tienen las curvas de indiferencia debe traducirse en supuestos acerca de la relación $\succsim$.

Estos supuestos son de dos clases: los de la forma misma y los de la dirección de mejora. Los de la forma son postulados que dices que los agentes prefiers canastas "balanceadas" a canastas "desbalanceadas," mientras que los de dirección de mejora dicen que los agentes prefieren más a menos. En términos de $\succsim$, estos supuestos se pueden hacer de diferentes formas. Sobre la primera propiedad, hay dos versiones:

Definición 1 $\succsim$ *es convexa si* $\forall x, x' \in \mathbb{R}^L_+$ *tales que* $x \succsim x'$ *y* $\forall \theta \in [0,1]$, $\theta x + (1-\theta) x' \succsim x'$.

Definición 2 $\succsim$ *es estríctamente convexa si* $\forall x, x' \in \mathbb{R}^L_+$ *tales que* $x \succsim x'$ *y* $x \neq x'$ *y* $\forall \theta \in (0,1)$, $\theta x + (1-\theta)x' \succ x'$.

Convexidad le da a las curvas de indiferencia su forma habitual, pero permite que haya trozos rectos en ellas. Convexidad estricta elimina esta posibilidad.

Ejercicio 2 *Demuestre que si* $\succsim$ *es estríctamente convexa, entonces es convexa.*

Ahora, para poder estudiar la propiedad de que más es mejor (no saciabilidad), necesitamos definir qué quiere decir "más" en $\mathbb{R}^L$, lo cual puede no ser obvio cuando $L > 1$.

Notación 1 *Para* $x = (x_1, ..., x_L), x' = (x'_1, ..., x'_L) \in \mathbb{R}^L$, *decimos que:*

- $x \geqslant x'$ *si* $\forall l \in \{1, ..., L\}$ $x_l \geqslant x'_l$.
- $x > x'$ *si* $x \geqslant x'$ *y* $x \neq x'$.
- $x \gg x'$ *si* $\forall l \in \{1, ..., L\}$ $x_l > x'_l$.

Con esto,

Definición 3 $\succsim$ *es monótona si* $\forall x, x' \in \mathbb{R}^L_+$, $x \gg x'$ *implica que* $x \succ x'$.

Definición 4 $\succsim$ *es estríctamente monótona si* $\forall x, x' \in \mathbb{R}^L_+$, $x > x'$ *implica que* $x \succ x'$.

Monotonicidad implica que en cualquier punto, el ortante abierto superior a él está estrictamente por encima de la curva de indiferencia. Monotonicidad estricta implica que esto mismo ocurra para el ortante cerrado.

Ejercicio 3 *Demuestre que si* $\succsim$ *es estríctamente monótona, entonces es monotona.*

Ahora, los economistas solemos utilizar un objeto artificial que denominamos la función de utilidad. Aunque este artificio no es para nada necesario en la construcción de nuestra teoría del equilibrio general, su uso se ha popularizado pues permite una notación más compacta y porque suele ser útil al momento de aplicar la teoría, así sea en ejemplos simples. Vamos a seguir aquí esa convención aunque, hay que decirlo, al hacerlo perdemos algo de generalidad. Por simplicidad, vamos a suponer que:

Supuesto 2 $\succsim$ *es representable: existe una función contínua* $u : \mathbb{R}^L_+ \longrightarrow \mathbb{R}$ *tal que* $u(x) \geqslant u(x')$ *si, y sólo si,* $x \succsim x'$. *En este caso, se dice que* u *representa a* $\succsim$.

En realidad, lo mejor que uno podría hacer es un supuesto en $\succsim$, que sea suficiente para implicar que $\succsim$ es representable. Este supuesto es conocido en la literatura como continuidad: que la relación $\succsim$ sea preservada en límites (o lo que es igual, que los conjuntos en y por encima de las curvas de indiferencia sean cerrados)

Ejercicio 4 *Demuestre lo siguiente: si* u *representa a* $\succsim$, *entonces*

1. $x \succ x'$ *si, y sólo si,* $u(x) > u(x')$.
2. $x \sim x'$ *si, y sólo si,* $u(x) = u(x')$.

Ejercicio 5 *Demuestre lo siguiente: si* u *representa a* $\succsim$, *entonces*

1. *Si* $\succsim$ *es convexa, entonces* u *es cuasicóncava.*

2. *Si $\succsim$ es estrictamente convexa, entonces u es estrictamente cuasicóncava.*

3. *Si $\succsim$ es monótona, entonces u es monótona creciente.*

4. *Si $\succsim$ es estrictamente monótona, entonces u es estrictamente monótona creciente.*

5. *Si $f : \mathbb{R} \longrightarrow \mathbb{R}$ es estrictamente creciente, entonces $v : f \circ u$ representa a $\succsim$. (Muestre además que es necesario que f sea estrictamente creciente.)*

Nos queda, por supuesto, estudiar el objeto $\mathcal{B}$. Como es habitual, definimos, para todo $p \in \mathbb{R}^L_{++}$ y todo $m \in \mathbb{R}_+$ el conjunto presupuestal:

$$B(p,m) = \left\{ x \in \mathbb{R}^L_+ \middle| \, p \cdot x \leqslant m \right\}$$

Con esto, la familia $\mathcal{B}$ se define como el conjunto de todos los posibles conjuntos presupuestales:

$$\mathcal{B} = \left\{ B \subseteq \mathbb{R}^L \middle| \, \exists (p,m) \in \mathbb{R}^L_{++} \times \mathbb{R}_+ : B(p,m) = B \right\}$$

Suponemos que cuando el agente enfrenta un conjunto presupuestal definido por precios p e ingreso m, escoge un elemento que es máximo con respecto a sus preferencias $\succsim$. Cuando las preferencias son representables, este último hecho es muy fácil de representar: decimos que el agente resuelve el siguiente problema:

$$\max_x u(x) \quad \text{s.a.} \quad x \in B(p,m)$$

Teorema 1 *Cuando las preferencias son representables, existe una solución al anterior problema.*

Prueba. Por definición, la función u que representa las preferencias es contínua. Como $m \geqslant 0$ y $p \gg (0,...,0)$, el conjunto $B(p,m)$ es compacto. Esto es suficiente para la existencia de un maximizador. ■

Cualquier x que solucione el anterior problema es una demanda óptima para el consumidor. Como la demanda es un elemento fundamental de nuestra teoría, en adelante siempre supondremos que $\succsim$ es representable, y trabajaremos directamente con una función de utilidad contínua u.

Dependiendo de qué otros supuestos hiciéramos en $\succsim$, tendríamos diferentes propiedades en u y, por ende, diferentes propiedades en la demanda:

Teorema 2 *Si u es cuasicóncava y $x, x' \in \mathbb{R}^L_+$ son soluciones al problema de maximización a precios p e ingreso m, entonces $\forall \theta \in [0,1]$, la canasta $\theta x + (1-\theta) x'$ también es solución al problema.*

Prueba. Queda como ejercicio. ■

Teorema 3 *Si u es estrictamente cuasicóncava, entonces para todo p y todo m, la solución al problema de maximización es única.*

Prueba. Queda como ejercicio. ■

Dado el anterior teorema, queda claro que cuando las preferencias son estrictamente convexas, uno puede definir una función de demanda $x : \mathbb{R}^L_{++} \times \mathbb{R}_+ \longrightarrow \mathbb{R}^L_+$. Cuándo esta condición es violada, lo máximo que uno puede es establecer una correspondencia de demanda. El teorema 2 implica que tal correspondencia es de valores convexos.

Abora, el segundo tipo de propiedad implica la siguiente propiedad fundamental:

Teorema 4 (La ley de Walras) *Si u es monótona y x resuelve el problema de optimización a precios p e ingreso m, entonces $p \cdot x = m$.*

Prueba. Queda como ejercicio. ■

Y otro resultado que es de gran utilidad es que:

Teorema 5 *Dada u es estríctamente cuasicóncava, la función de demanda $x : \mathbb{R}^L_{++} \times \mathbb{R}_+ \longrightarrow \mathbb{R}^L_+$ es contínua.*

Prueba. La prueba es algo complicada: es una aplicación del teorema del máximo. ■

En realidad, la cuasiconcavidad estricta sólo se utiliza para garantizar que la función de demanda esté definida. Sin ese supuesto, la correspondencia de demanda aún mantiene una propiedad de continuidad que es muy util.

En adelante, mantendremos los supuestos de que las funciones de utilidad son estrictamente cuasicóncavas y estríctamente monótonas.

Bajo estas propiedades, el siguiente resultado es intuitivo:

Teorema 6 *La demanda satisface el axioma debil de las preferencias reveladas: si $p \cdot f(p', m') \leqslant m$ y $f(p, m) \neq f(p', m')$, entonces $p' \cdot f(p, m) > m'$.*

Prueba. Suponga que no: $p' \cdot f(p, m) \leqslant m'$. Como $p \cdot f(p', m') \leqslant m$, se sigue que $u(f(p, m)) \geqslant u(f(p', m'))$. Como $p' \cdot f(p, m) \leqslant m'$, se sigue que $p' \cdot \left(\frac{1}{2} f(p, m) + \frac{1}{2} f(p', m')\right) \leqslant m'$, pero por cuasiconcavidad estricta, $u\left(\frac{1}{2} f(p, m) + \frac{1}{2} f(p', m')\right) > u(f(p', m'))$, contradiciendo el hecho de que $f(p', m')$ es la demanda óptima dado (p', m'). ■

Notas de clase sobre Equilibrio General.
Nota 3: Economías de Intercambio.

Andrés Carvajal
Universidad del Rosario, Departamento de Economía. acarvaes@banrep.gov.co

October 14, 2003

Abstract

En el trabajo que se ha hecho en los dos cursos anteriores de la secuencia de micro, el análisis se ha enfocado o bien en el problema de un agente individualmente considerado, o bien en el mercado de un bien, individualmente considerado. Este enfoque puede ser conveniente en algunos casos, pero ignora fenómenos de gran importancia, en particular el de cómo lo que ocurre en un mercado afecta el comportamiento, deseado o posible, de los individuos en los demás mercados. No me refiero al efecto obvio que se presenta cuando los mercados en cuestión son los de bienes substitutos o complementarios: dichos efectos deben haber sido estudiados en cursos previos. El punto central es que ahora permitamos que el presupuesto que los agentes pueden gastar no sea independiente de los precios.

Lo que vamos a intentar hacer por medio de la teoría del EG es ambicioso desde tres puntos de vista:

- Vamos a tratar de analizar simultáneamente la totalidad de mercados de bienes y la determinación de sus precios de equilibrio.
- Vamos, para ello, a endogenizar variables que hasta ahora han sido consideradas como exógenas, en particular el ingreso de los agentes.
- Vamos a intentar todo esto sin utilizar nuevos supuestos, más allá de los que se han utilizado hasta ahora en los cursos de microeconomía:
  - Cada agente es descrito por sus preferencias.
  - Cada firma es descrita por su tecnología.
  - Institucionalmente, para cada bien existe un precio que aplica para todos los consumidores y firmas de la economía.
  - Cada consumidor maximiza su bienestar por medio de su demanda, tomando los precios como dados y sujeto únicamente a no gastar más que el valor de su riqueza.
  - Cada firma maximiza sus ganancias tomando los precios como dados y sujeto únicamente a su teconología.

En este contexto, el concepto de equilibrio que vamos a utilizar es precisamente el de Walras: la economía se encuentra en equilibrio cuando los precios son tales que la suma de las demandas individuales óptimas es igual a la suma de las ofertas individuales y los bienes pre-existentes.

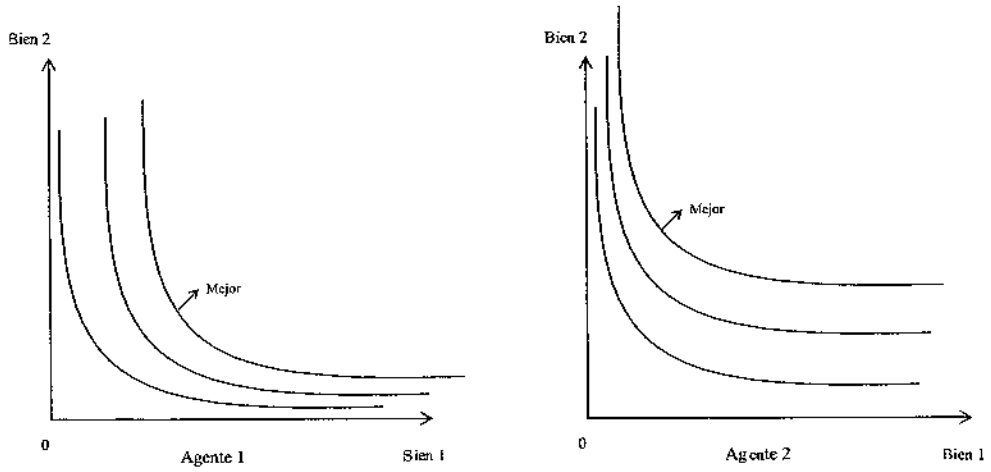
¿Por qué es esto equilibrio? Porque los agentes encuentran una situación en la que todos pueden hacer lo que consideran más conveniente para ellos mismos.

Como dije con anterioridad, las complicaciones que tenemos que resolver son principalmente generadas por los ingresos. Los efectos ingreso, como es bien sabido, son un problema de teoría del consumidor, y no de teoría del productor. Lo que quiero decir es que el problema fundamental de EG contiene los elementos más interesantes aún si nos abstraemos de aspectos de producción y consideramos solamente una economía de intercambio y de consumo. Esto es precisamente lo que vamos a hacer.

1 Una economía de 2×2

A manera introductoria suponga que hay una economía en la que sólo existen dos consumidores y dos bienes, y en la que los supuestos de competencia perfecta son satisfechos.

Denominaremos a los agentes 1 y 2. En general, nos referiremos a los agentes con el super-índice i. Denominaremos a los bienes 1 y 2. En general, nos referiremos a los bienes por el sub-índice l. Supongamos que los mapas de indiferencia que caracterizan a cada agente son como a continuación:

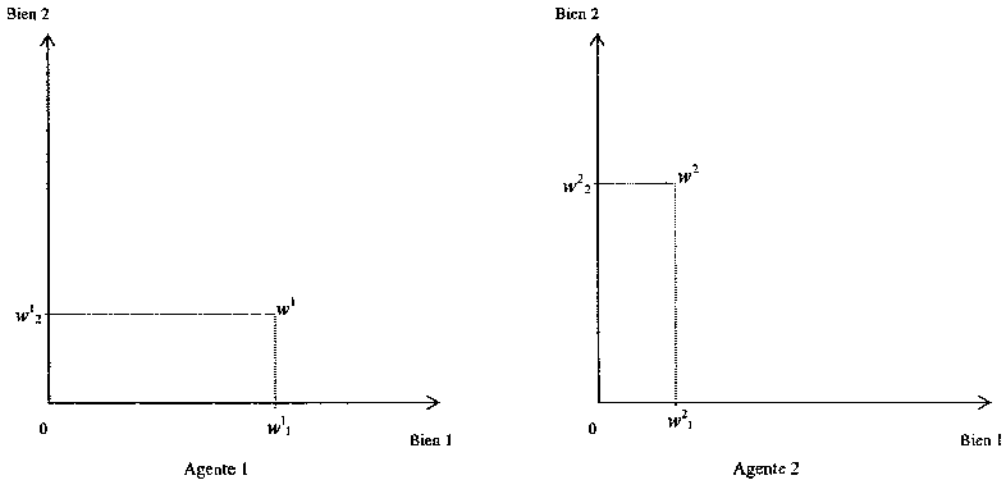


Supondremos que estas curvas de indiferencia están respectivamente asociadas a funciones de utilidad u^1 y u^2, cada una de las cuales tiene como dominio el cuadrante no negativo de $\mathbb{R}^2$ y como codominio a $\mathbb{R}$.

Ahora, como estamos suponiendo que ésta es una economía en la que los bienes no se producen sino que solamente se intercambian, vamos a suponer que cada agente posee una cantidad no negativa de cada uno de los bienes.

Denotaremos por w_j^i la cantidad que el agente i posee del bien j. Por ejemplo, $w_1^2 = 5$ quiere decir que el agente 2 posee 5 unidades del bien 1.

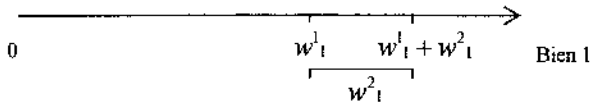
La dotación del agente i es el par (w_1^i, w_2^i) el cual denotaremos por w^i. (Así, $w^2 = (3, 5)$ significa que el agente 2 tiene, respectivamente, 3 y 5 unidades de los bienes 1 y 2.) Con esto, un segundo componente de la caracterización de cada agente es su dotación:



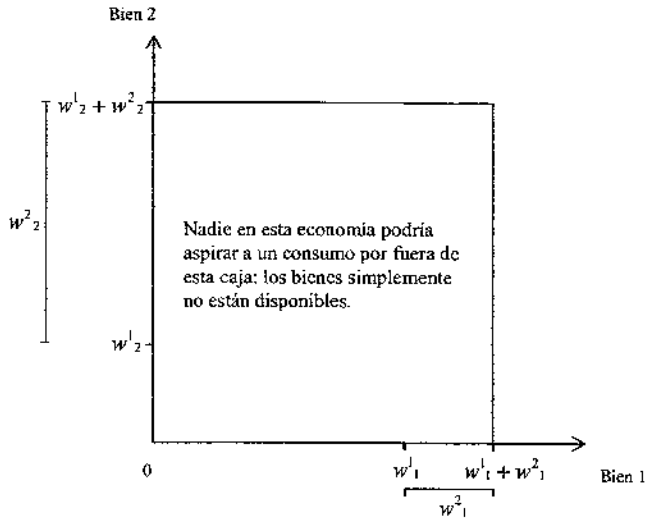
De igual forma, dado que los agentes no pueden obtener más bien que su dotación, la cantidad que la economía posee de cada bien está determinada por el agregado de las dotaciones individuales.

1.1 La caja de Edgeworth:

Consideremos inicialmente el bien 1. Sobre la linea real, la dotación del agente 1 y la dotación agregada son respectivamente los puntos w_1^1 y $\left(w_1^1 + w_1^2\right)$. La distancia entre estos dos puntos equivale a la dotación del agente 2, w_1^2. Gráficamente:

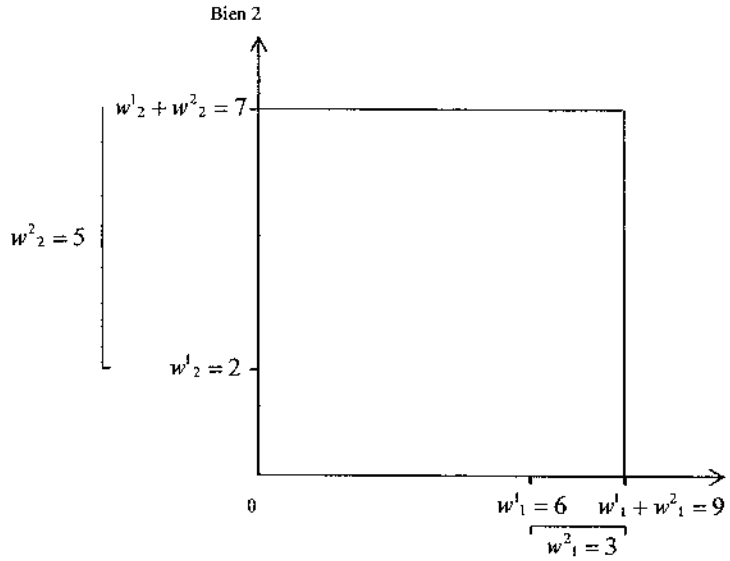


Sobre esta última gráfica, hagamos el mismo análisis para el bien 2:



Así, completando los lados de la caja obtenemos el conjunto total de valores de los bienes que la economía puede factiblemente distribuir.

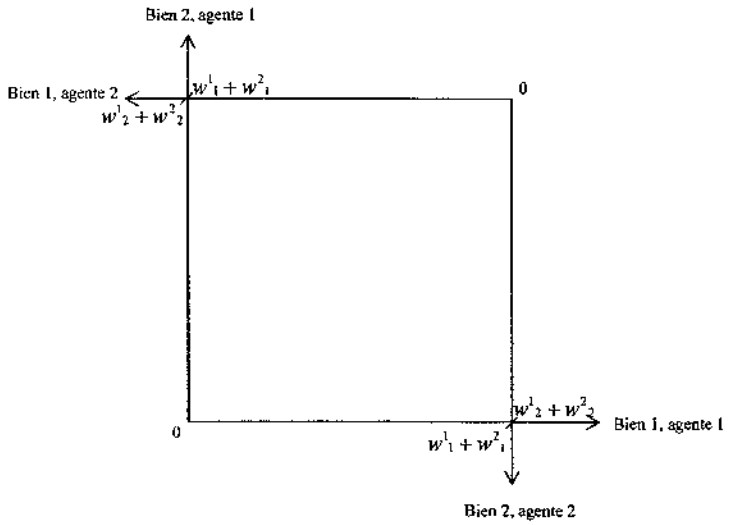
Ejemplo 1 *Si* $w^1 = (6,2)$ *y* $w^2 = (3,5)$, *entonces el total disponible del bien* 1 *es* 9 *y del bien* 2 *es* 7. *Gráficamente:*



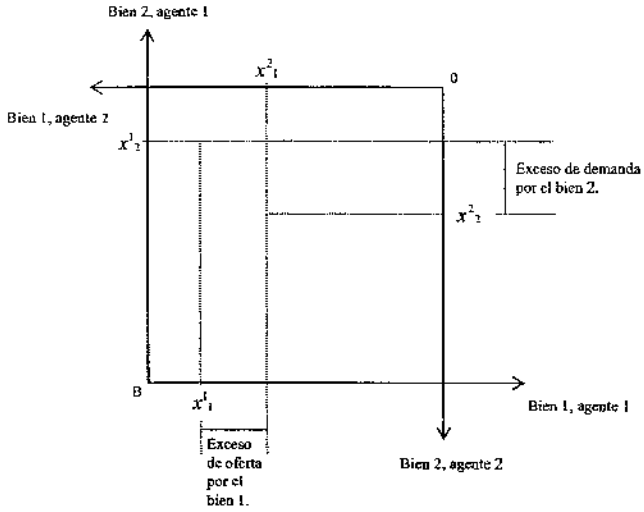
Denotaremos por x_l^i el consumo del bien l por el agente i (es decir, $x_2^1 = 1$ quiere decir que el agente 1 consumió 1 unidades del bien 2).

Denotaremos por x^i el consumo del agente i. Es decir: $x^i = \left(x_1^i, x_2^i\right)$. Así, $x^1 = (1,6)$ quiere decir que el agente 1 consumió, respectivamente, 1 y 6 unidades de los bienes 1 y 2.

Podemos utilizar los lados superior y derecho de la caja para medir los consumos del agente 2, midiendo en el lado superior (y hacia la izquierda) el consumo del bien 1 y en el lado derecho (y hacia abajo) su consumo del bien 2.

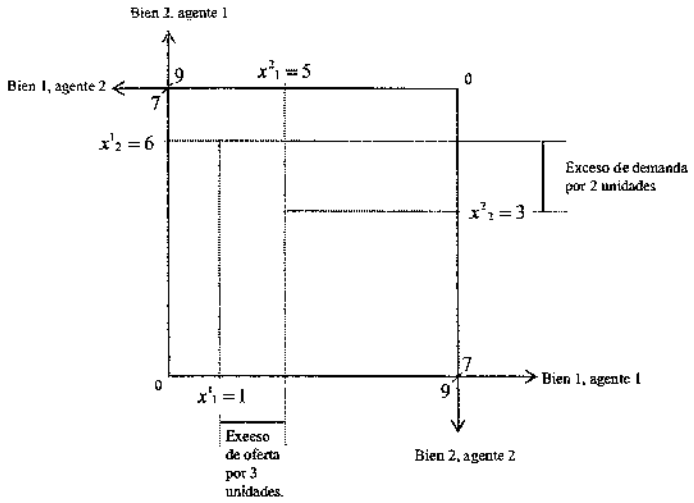


Con esta convención, dados consumos $x^1 = (x^1_1, x^1_2)$ y $x^2 = (x^2_1, x^2_2)$ podemos estudiar los excesos de oferta/demanda simultáneamente:

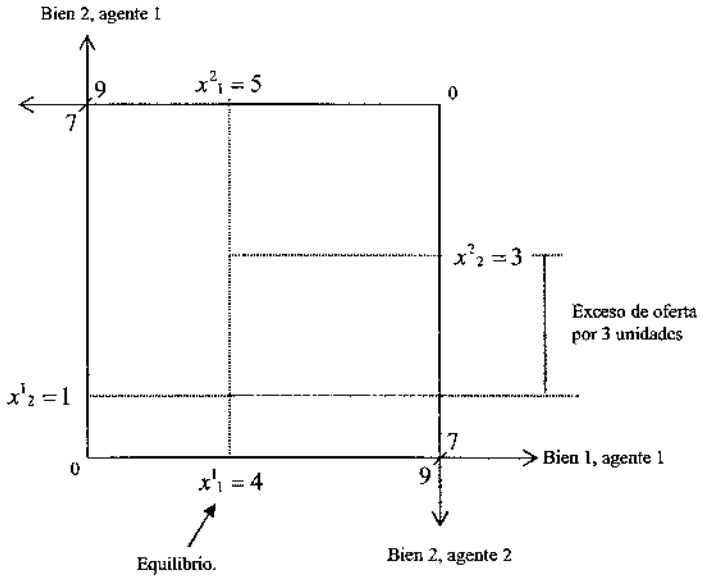


Ejemplo 2 *Supongamos nuevamente que* $w^1 = (6,2)$ *y* $w^2 = (3,5)$. *Supongamos que* $x^1 = (1,6)$ *y* $x^2 = (5,3)$. *En este caso, la demanda total por el bien* 1 *es* 7 *y por el bien* 2 *es* 9. *Esto crea un exceso de oferta*

de bien 1 por 3 unidades y un exceso de demanda por bien 2 de 2 unidades. Gráficamente:



Si, alternativamente, suponemos que $x^1 = (4,1)$ *y* $x^2 = (5,3)$*, entonces la oferta y la demanda por bien 1 son iguales, pero hay ahora un exceso de oferta de bien 2 por 3 unidades. Esto es:*



Ejercicio 1 *Con las mismas dotaciones del ejemplo anterior, determine algebráica y gráficamente los excesos de oferta/demanda correspondientes a los siguientes pares de consumos:*

1. $x^1 = (2,5)$ *y* $x^2 = (4,3)$
2. $x^1 = (4,1)$ *y* $x^2 = (6,5)$
3. $x^1 = (1,1)$ *y* $x^2 = (3,3)$

4. $x^1 = (5,5)$ *y* $x^2 = (5,4)$

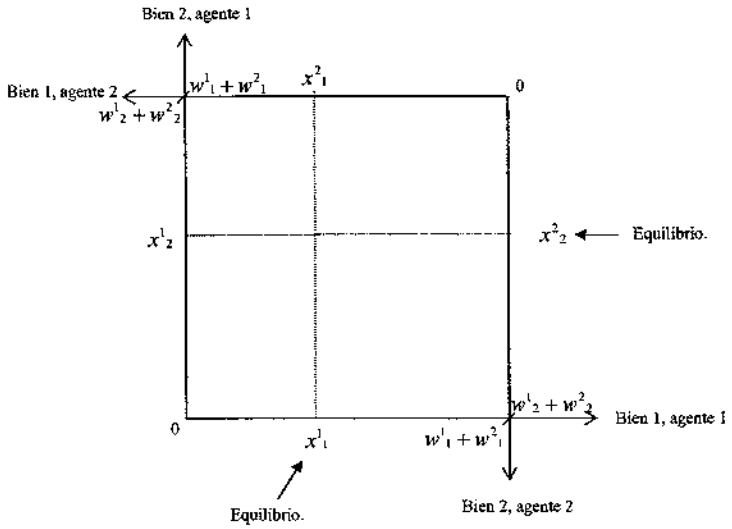
5. $x^1 = (0,7)$ *y* $x^2 = (1,1)$

6. $x^1 = (0,7)$ *y* $x^2 = (9,0)$

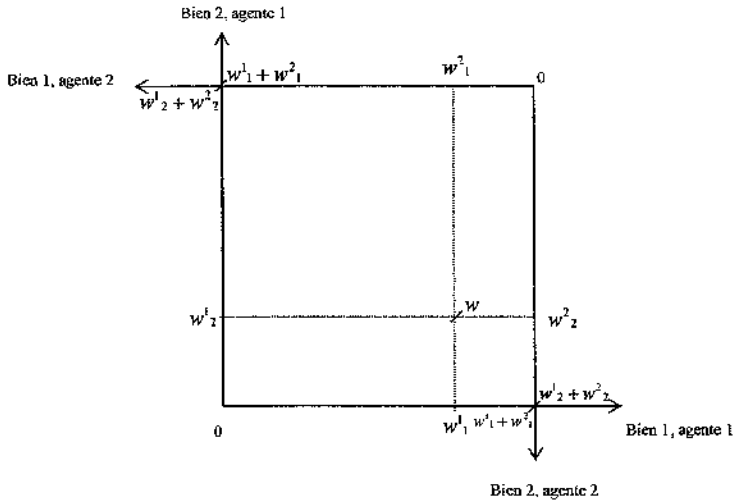
7. $x^1 = (4,4)$ *y* $x^2 = (5,3)$

8. $x^1 = w^1$ *y* $x^2 = w^2$

Por supuesto, existen pares de consumos ante los cuales no se generan excesos de oferta o demanda:

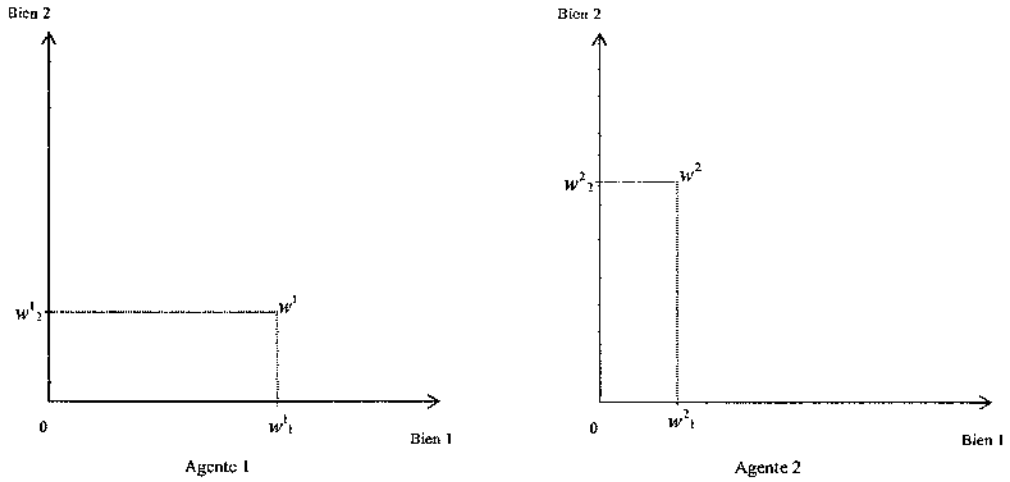


Este tipo de pares de consumos que no generan excesos de oferta o demanda se denominan factibles. Un caso particular de este tipo de factibilidad ocurre cuando $x^1 = w^1$ y $x^2 = w^2$:

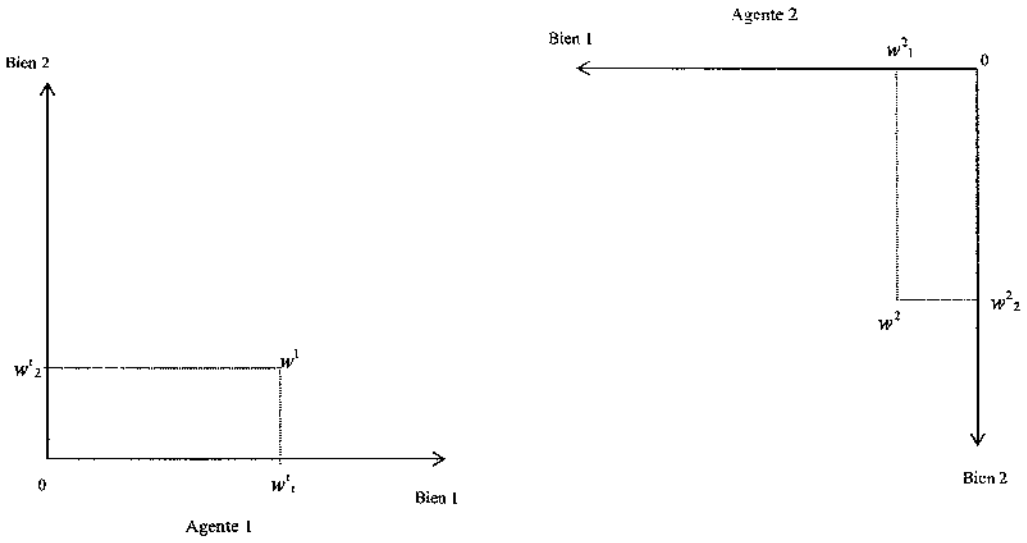


Este último hecho, que es algebráicamente obvio, lo hubieramos podido encontrar por otro método algo más mecánico:

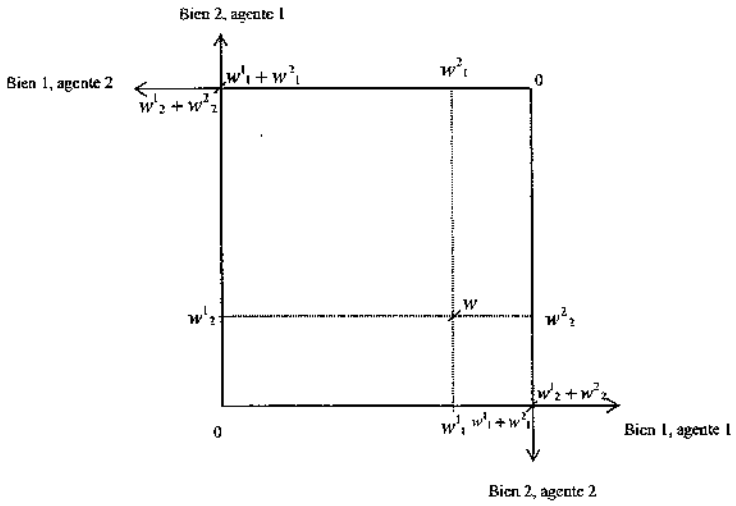
1. Tomemos las gráficas de las dotaciones de cada uno de los agentes:



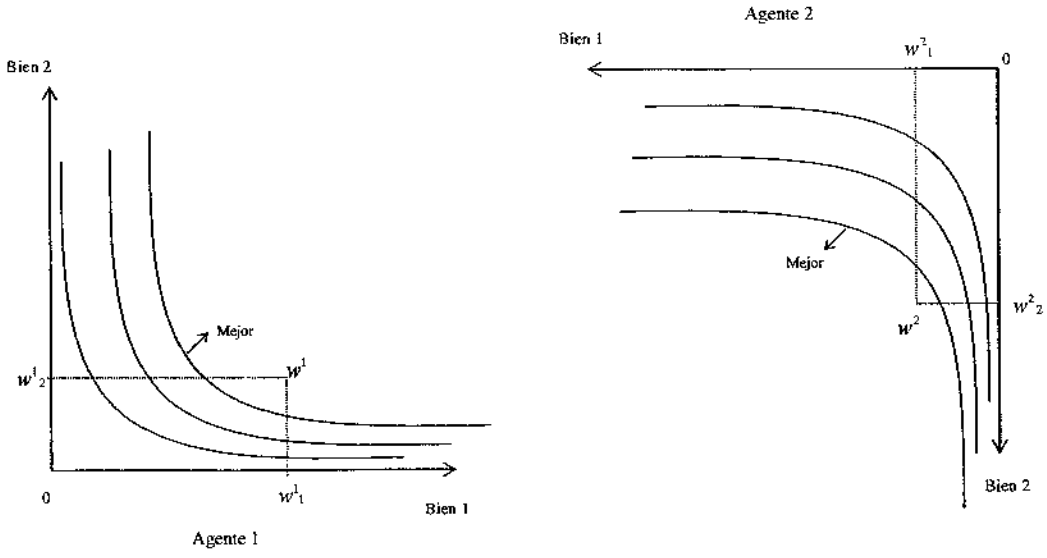
2. Rotemos 180 grados la gráfica correspondiente al agente 2:



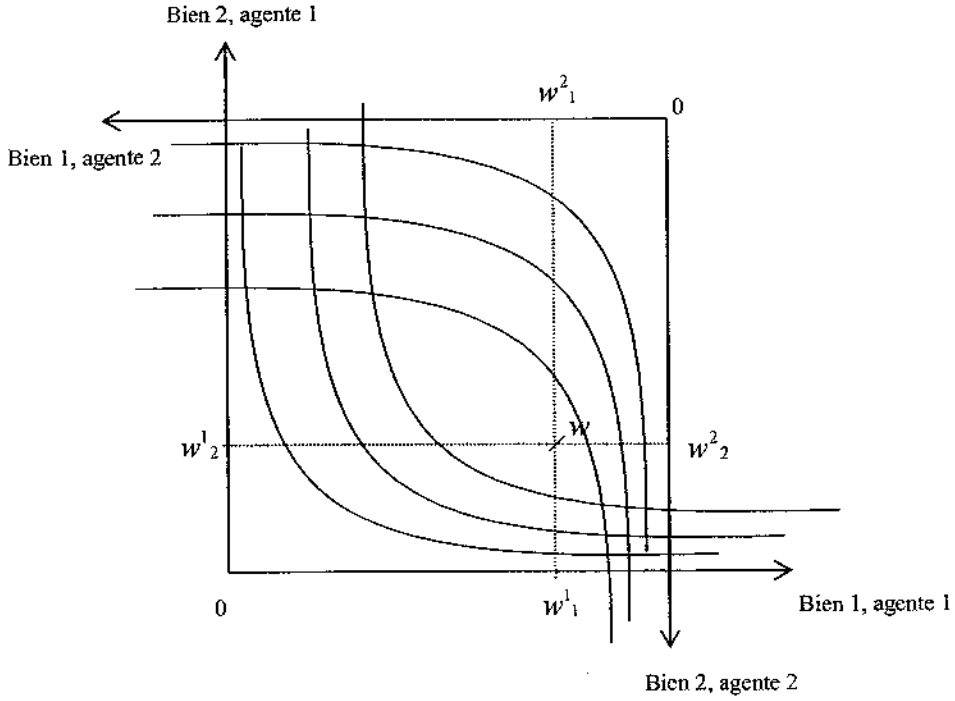
3. Y ahora empalmemos las dos gráficas de forma tal que los puntos w^1 y w^2 coincidan:



Este mecanismo de rotar el espacio de bienes del agente 2 es muy importante por la siguiente razón: recuerde que asociado al espacio de consumo viene el mapa de indiferencia; entonces, al rotar el plano para el agente 2, al mismo tiempo estamos rotando su mapa de indiferencias, de forma tal que, implícito en el segundo paso tenemos:



Y por lo tanto, al empalmar las dos gráficas:



Esta última gráfica se conoce como la Caja de Edgeworth de la economía y es el instrumento básico para el análisis gráfico del equilibrio general de una economía de intercambio 2×2.

1.2 El análisis de Walras:

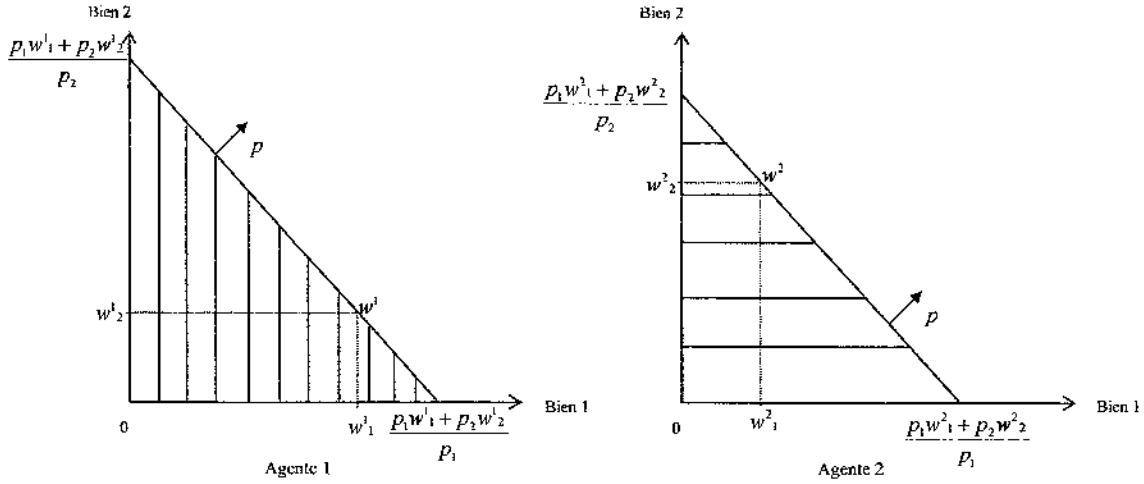
1.2.1 Los precios y las restricciones presupuestales:

Como hemos dicho antes, Walras centró su análisis en el intercambio de bienes por medio de mercados en los que para todos los bienes está definido un precio al cual todos los agentes pueden comprar o vender el bien.

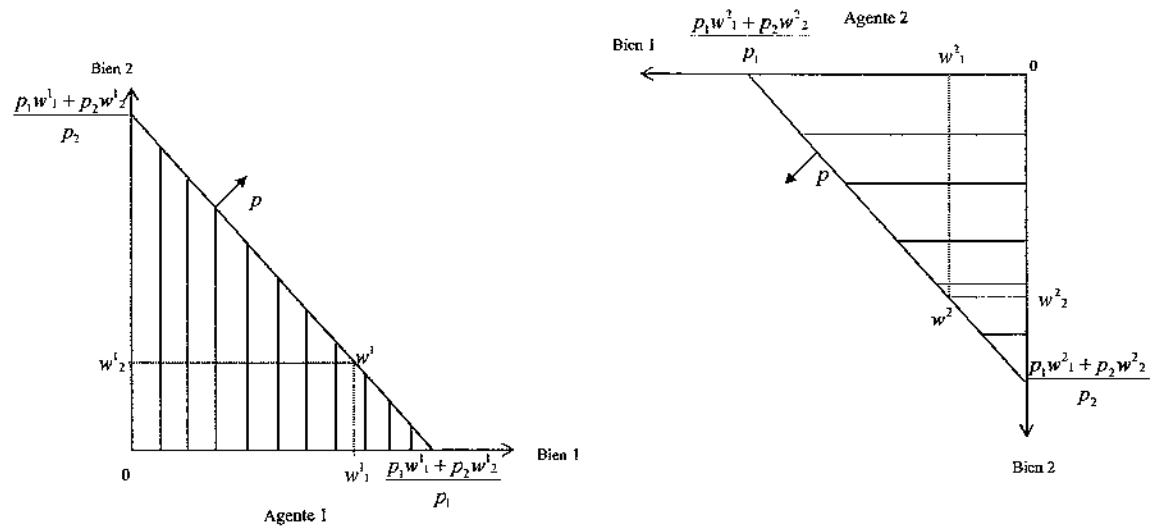
Denotaremos el precio del bien l por p_l. Así, el par de precios p es (p_1, p_2). Supondremos que los precios son estrictamente positivos.

Como en teoría del consumidor, en la economía competitiva suponemos que cada consumidor toma los precios como dados y determina su demanda de forma tal que maximice su bienestar sujeto únicamente a la restricción de que el valor de su consumo a los precios del mercado no supere el valor de su dotación (su ingreso) a esos mismos precios.

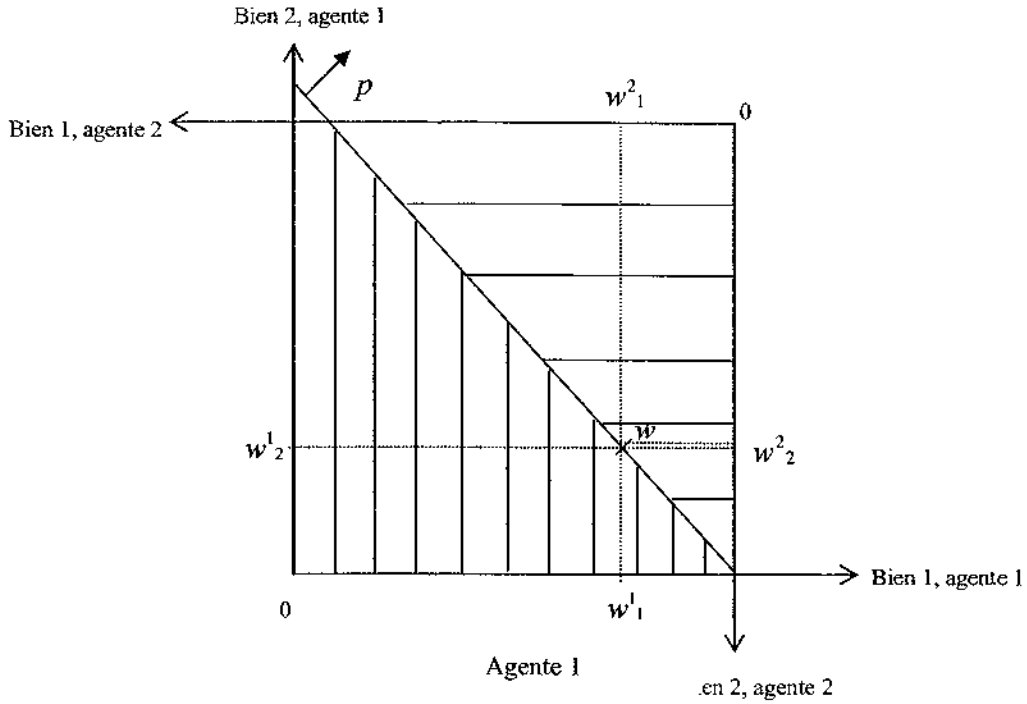
Gráficamente, a precios p:



Donde el area con líneas verticales en el panel de la izquierda muestra el conjunto presupuestal (demandas factibles) a precios p para el consumidor 1 y el area con líneas horizontales en el panel de la derecha muestra lo correspondiente al consumidor 2. Como con las curvas de indiferencia, cuando rotamos el espacio de bienes del agente 2, su conjunto presupuestal debe también rotarse:



Con esto, al empalmar las gráficas como en la caja de Edgeworth obtenemos:

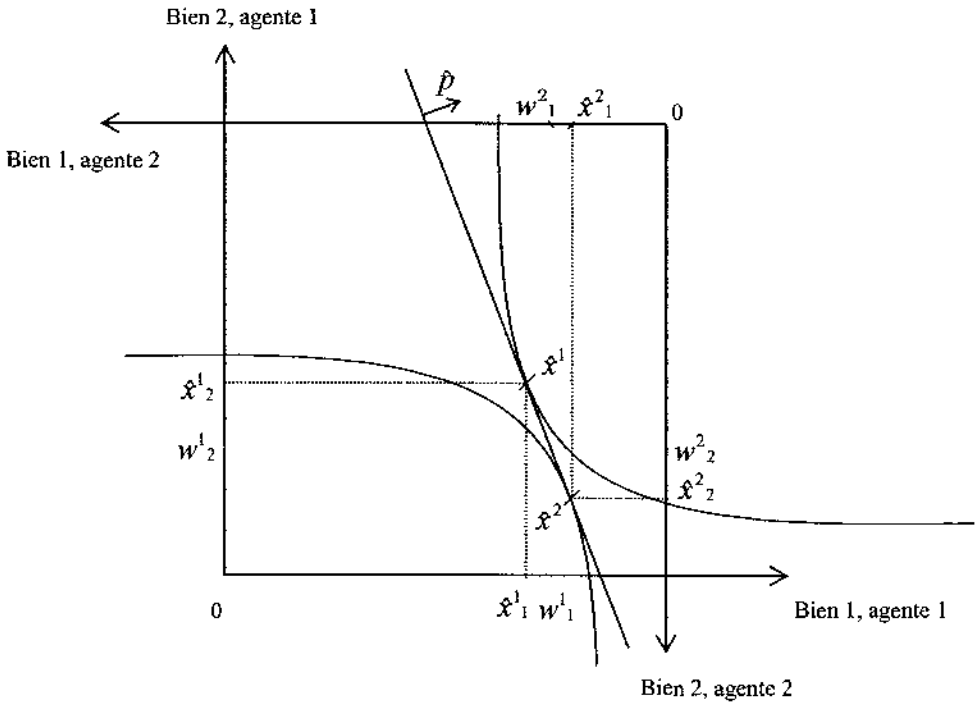


donde, nuevamente, el area con líneas verticales es el conjunto presupuestal para el consumidor 1 (incluído el triángulo que sale en la esquina superior izquierda de la caja) y el area con líneas horizontales es el conjunto presupuestal del consumidor 2.

1.2.2 El Equilibrio Walrasiano:

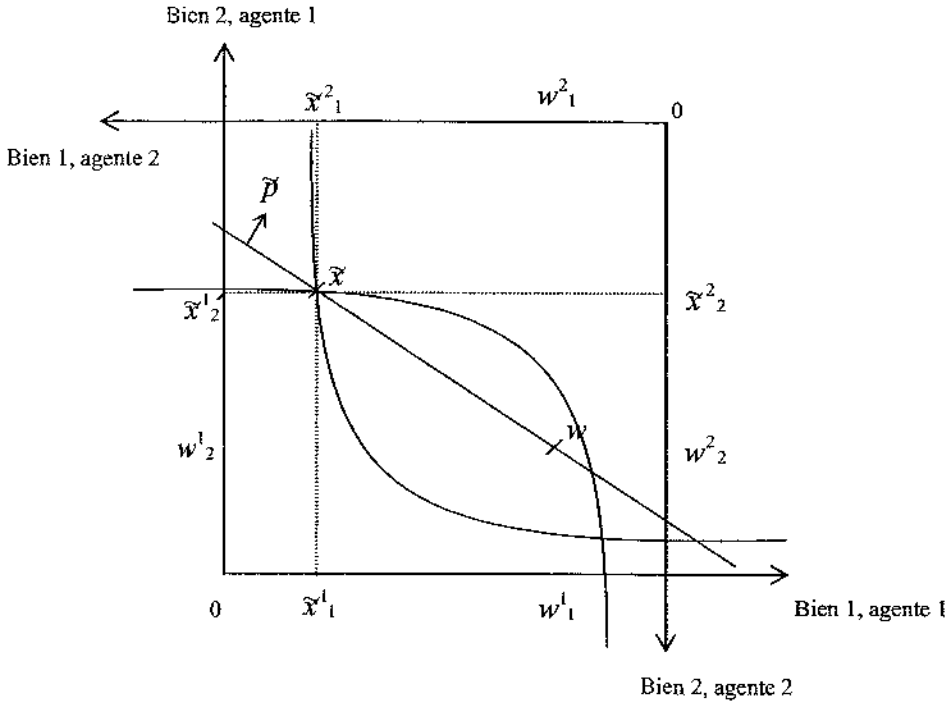
Consideremos tres casos diferentes:

Caso 1 *El vector de precios es $\widehat{p}$ y las demandas son $\widehat{x}^1$ y $\widehat{x}^2$, como en la siguiente gráfica:*



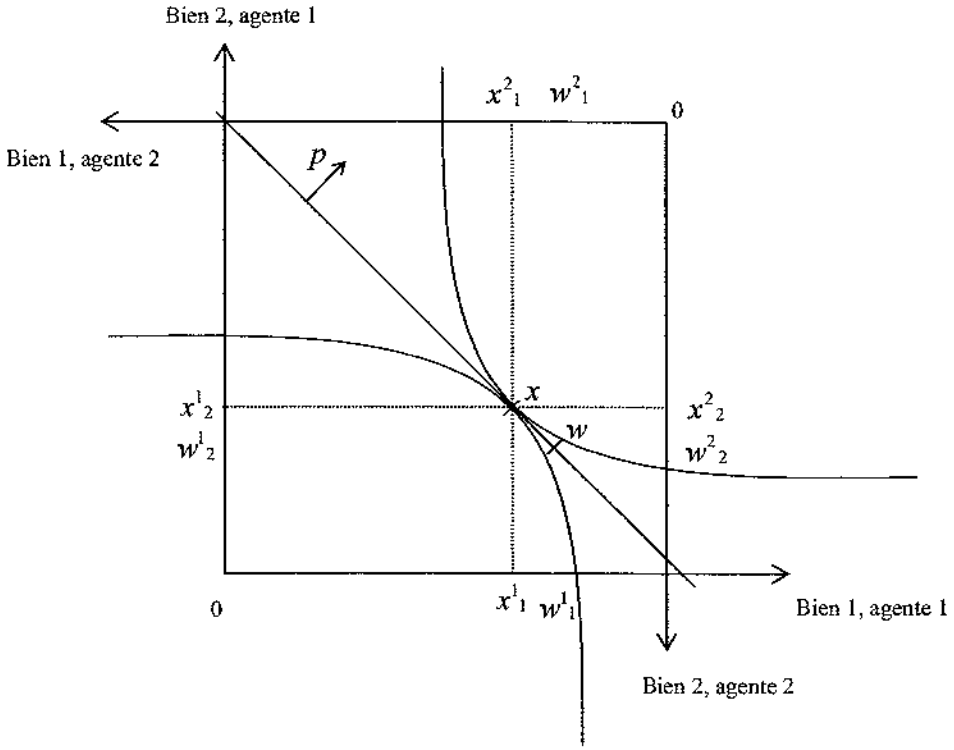
En este caso, a los precios $\widehat{p}$ la canasta $\widehat{x}^1$ maximiza el bienestar del agente 1, sujeto a que el valor de su consumo no supere el valos de su dotación. Similarmente, la canasta $\widehat{x}^2$ maximiza el bienestar del agente 2, sujeto a que el valor de su consumo no sobrepase el de su dotación. Esta situación, sin embargo, no es de equilibrio: hay exceso de oferta del bien 1 y de demanda del bien 2 (es decir, la situación no es factible).

Caso 2 *El vector de precios es $\widetilde{p}$ y las demandas son $\widetilde{x}^1$ y $\widetilde{x}^2$, como en la siguiente gráfica:*



En esta situación, las canastas $\widetilde{x}^1$ y $\widetilde{x}^2$ no generan excesos de oferta o demanda en los mercados. Sin embargo, la situación no es de equilibrio, pues a los precios $\widetilde{p}$ la canasta $\widetilde{x}^1$ no maximiza el bienestar del agente 1 sujeto a la restricción de que el valor de su consumo no supere al de su dotación, ni la canasta $\widetilde{x}^2$ lo hace con el bienestar del agente 2.

Caso 3 *El vector de precios es p y las demandas son x^1 y x^2, como en la siguiente gráfica:*



En este caso, la canasta x^1 maximiza el bienestar del agente 1 sujeto a que el valor de su consumo no supere al de su dotación. La canasta x^2 hace lo mismo con el bienestar del agente 2. Adicionalmente, los mercados se equilibran, pues no se generan excesos de oferta o demanda.

Esta es una situación de equilibrio general como la concibió Walras: cada agente (*tomando los precios como dados*) se encuentra feliz con el resultado del intercambio (pues a esos precios y con su dotación inicial el agente no podría estar mejor) y además el intercambio puede realizarse (pues nadie está tratando de vender o comprar sin poder hacerlo).

Formalmente, en nuestra economía de 2×2 :

Definición 1 *Dadas las preferencias de los agentes (u^1 y u^2) y sus dotaciones (w^1 y w^2) un equilibrio general es un par compuesto por unos precios p y unas demandas individuales x^1 y x^2 tales que:*

1. *Para cada agente, su demanda maximiza su bienestar sujeto a su restricción presupuestal:*

   Para el agente 1:

$$u^1\left(x_1^1, x_2^1\right) = \max u^1\left(x_1, x_2\right) \quad \textit{sujeto a: } p_1x_1 + p_2x_2 \leqslant p_1w_1^1 + p_2w_2^1$$

   Para el agente 2:

$$u^2\left(x_1^2, x_2^2\right) = \max u^2\left(x_1, x_2\right) \quad \textit{sujeto a: } p_1x_1 + p_2x_2 \leqslant p_1w_1^2 + p_2w_2^2$$

2. *Los mercados están en equilibrio:*

Para el bien 1:

$$x_1^1 + x_1^2 = w_1^1 + w_1^2$$

Para el bien 2:

$$x_2^1 + x_2^2 = w_2^1 + w_2^2$$

Ejemplo 3 *Supongamos que*

$$\begin{aligned} u^1(x_1, x_2) &= \sqrt{x_1 x_2} \\ u^2(x_1, x_2) &= \sqrt{x_1 x_2} \\ w^1 &= (1, 1) \\ w^2 &= (1, 1) \end{aligned}$$

Dados estos datos, el problema para el consumidor 1 es:

$$\max \sqrt{x_1 x_2} \quad \textit{sujeto a: } p_1 x_1 + p_2 x_2 \leqslant p_1 + p_2$$

cuya solución es

$$\begin{aligned} x_1^1(p_1, p_2) &= \frac{1}{2}\frac{p_1 + p_2}{p_1} \\ x_2^1(p_1, p_2) &= \frac{1}{2}\frac{p_1 + p_2}{p_2} \end{aligned}$$

Dado que el problema de 2 es idéntico:

$$\begin{aligned} x_1^2(p_1, p_2) &= \frac{1}{2}\frac{p_1 + p_2}{p_1} \\ x_2^2(p_1, p_2) &= \frac{1}{2}\frac{p_1 + p_2}{p_2} \end{aligned}$$

Ahora busquemos precios p_1 y p_2 tales que los mercados se equilibren con estas demandas:

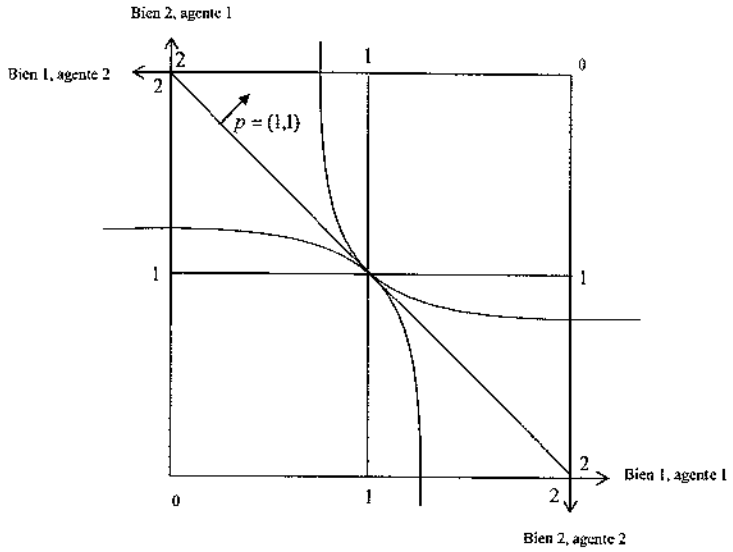
$$\begin{aligned} x_1^1(p_1, p_2) + x_1^2(p_1, p_2) &= w_1^1 + w_1^2 \\ x_2^1(p_1, p_2) + x_2^2(p_1, p_2) &= w_2^1 + w_2^2 \end{aligned}$$

Substituyendo:

$$\begin{aligned} \frac{1}{2}\frac{p_1 + p_2}{p_1} + \frac{1}{2}\frac{p_1 + p_2}{p_1} &= 1 + 1 \\ \frac{1}{2}\frac{p_1 + p_2}{p_2} + \frac{1}{2}\frac{p_1 + p_2}{p_2} &= 1 + 1 \end{aligned}$$

Este sistema tiene infinitas soluciones. Cualquier par de p_1 y p_2 que satisfaga que $p_1 > 0$ y $p_1 = p_2$ soluciona el sistema. Por tanto, $p = (1, 1)$, $x^1 = x^1(1, 1) = (1, 1)$ y $x^2 = x^2(1, 1) = (1, 1)$ es un equilibrio general.

Gráficamente:



Nota técnica 1 *Cuando la restricción presupuestal de un agente está dada por el valor de una dotación,su demanda no cambia si uno multiplica todos los precios por una constante positiva. Por esta razón, en cualquier ecconomía hay un número infinito de vectores de precios de equilibrio: si $p = (p_1, p_2)$ es un vector de precios de equilibrio, también lo son $(2p_1, 2p_2)$, $\left(\frac{1}{9}p_1, \frac{1}{9}p_2\right)$, $(500p_1, 500p_2)$ y, en general, cualquier producto de p por un número positivo. Por esta razón, uno suele "normalizar" los precios fijando, por ejemplo, $p_1 = 1$ o requiriendo que $p_1 + p_2 = 1$.*

Ejercicio 2 *Encuentre los equilibrios de las siguientes economías:*

1. $u^1(x_1, x_2) = \sqrt{x_1 x_2}$, $u^2(x_1, x_2) = \sqrt{x_1 x_2}$, $w^1 = (2, 1)$ *y* $w^2 = (1, 1)$.
2. $u^1(x_1, x_2) = x_1^{0.6} x_2^{0.6}$, $u^2(x_1, x_2) = x_1^{0.4} x_2^{0.4}$, $w^1 = (2, 2)$ *y* $w^2 = (0, 0)$.
3. *(Más difícil) Para* $0 < a < 1$ *y* $0 < b < 1$, $u^1(x_1, x_2) = x_1^a x_2^{1-a}$, $u^2(x_1, x_2) = x_1^b x_2^{1-b}$, w^1 *y* w^2 *arbitrarios.*

1.2.3 La ley de Walras:

Retomemos el ejemplo 3. Supongamos que sólo huhiéramos buscado los precios que ajustan la oferta y la demanda del bien1:

$$x_1^1(p_1, p_2) + x_1^2(p_1, p_2) = w_1^1 + w_1^2$$

Es decir que,

$$\frac{1}{2}\frac{p_1 + p_2}{p_1} + \frac{1}{2}\frac{p_1 + p_2}{p_1} = 1 + 1$$

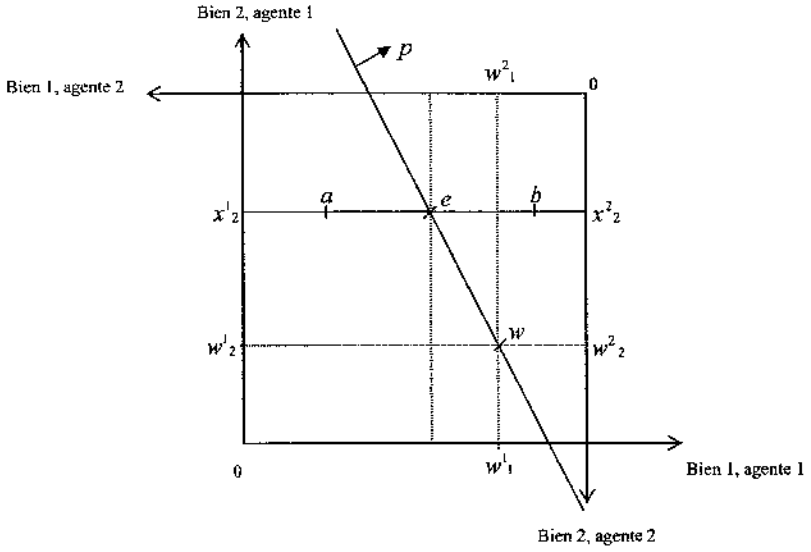
o, lo que es igual:

$$\frac{p_1 + p_2}{p_1} = 2$$

Si normalizamos $p_1 = 1$, es claro que $p_2 = 1$ soluciona el problema.

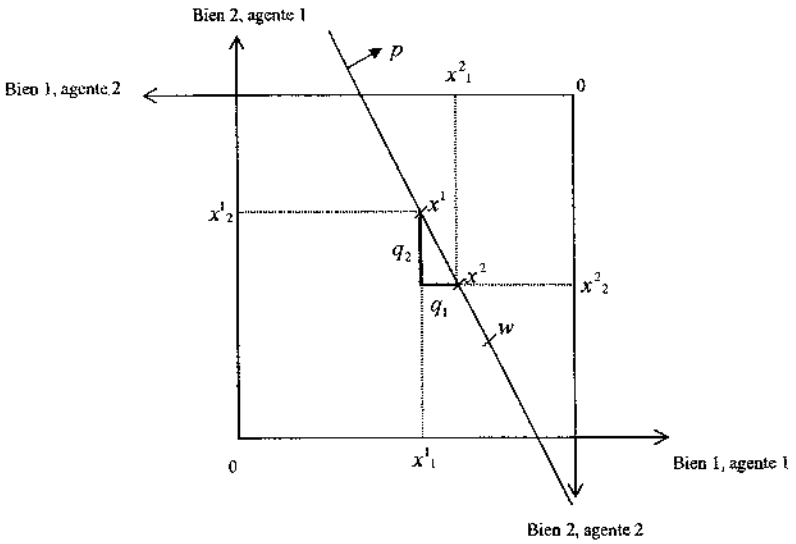
Lo importante que tenemos que observar es que ¡sólo considerando el bien 1, obtenemos los mismos precios de equilibrio que cuando consideramos ambos bienes! Es obvio que estos precios también equilibran el mercado del bien 2.

El punto es que en esta economía en la que los agentes toman sus decisiones maximizando su propio bienestar, si uno de los mercados se ha ajustado, también el otro, de manera automática debe haberlo hecho. Es fácil ver esto gráficamente, dados unos precios p:



En este caso, el mercado del bien 2 se ba ajustado. Como ya sabemos, para que el mercado del bien 1 no se encuentre en equilibrio, tendría que ocurrir que los puntos x^1 y x^2 no coincidan en la gráfica. Pero como el mercado del bien 2 se ha ajustado, tiene que ser que x^1 y x^2 se encuentran en la linea borizontal contínua que pasa por el punto e, pero siendo (al menos uno de ellos) diferente de e (por ejemplo, si $x^1 = a$ y $x^2 = b$, hay exceso de oferta del bien 1; si, por el contrario, $x^1 = b$ y $x^2 = a$, entonces el exceso es de demanda). Pero como estamos suponiendo que los agentes prefieren consumir más que consumir menos, entonces estos puntos, que están estrictamente por debajo de las restricciones presupuestales, no pueden ser óptimos individualmente. Puesto de otra forma, dados x^1_2 y x^2_2, el hecho de que los agentes escojan sus demandas de manera óptima implica que, gráficamente, $x^1 = e$ y $x^2 = e$ y que, por tanto, el bien 1 también se ajuste.

Otra forma de ver este hecho es la siguiente. Dados precios p, supongamos que x^1 y x^2 son como en el gráfico a continuación:



En este caso, q_1 es exceso de oferta de bien 1 y q_2 es exceso de demanda por bien 2. Ahora, dado que todos los puntos sobre la linea presupuestal del agente 1 cuestan lo mismo a precios p, debe ser que la suma de los valores de estos excesos es cero: adoptando la convención $q_1 < 0$ y $q_2 > 0$, tenemos que $p_1 q_1 + p_2 q_2 = 0$. De este resultado se sigue que, dado que p_1 y p_2 son positivos, si uno de los dos excesos se redujera a cero, lo mismo tendría que pasar con el otro.

Nota técnica 2 *Nótese que para que nuestros argumentos sean válidos necesitamos suponer que cada agente, al escoger óptimamente su canasta de consumo, siempre escoge un punto en la linea presupuestal. Para esto es suficiente suponer monotonicidad.*

Lo que hemos hecho es argumentar lo siguiente:

Teorema 1 *Si los agentes de una economía 2×2 escogen sus demandas óptimamente, el equilibrio entre oferta y demanda en uno de los mercados implica el mismo equilibrio en el otro mercado.*

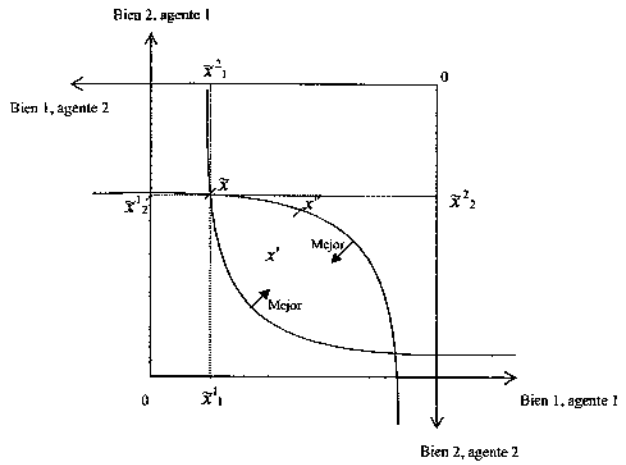
1.3 El análisis de Pareto:

1.3.1 Eficiencia:

Recordemos que Pareto tenía una concepción de lo que debía entenderse como equilibrio diferente a la que había propuesto Walras. Para él, los resultados del intercambio en los cuales la economía no encontraría impulsos para abandonarlos eran aquellos en los cuales la tensión entre lo que es posible y lo que es deseable es máxima en el sentido de que, dados los recursos de que dispone la economía, mejorar el bienestar de uno de los agentes implica empeorar el de algún otro.

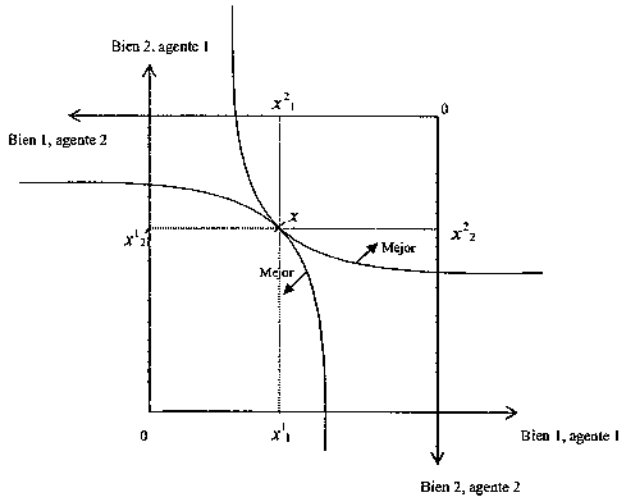
Utilizando nuevamente la caja de Edgeworth, consideremos los siguientes casos:

Caso 4 *Considere la asignación* $\widetilde{x}$*:*



La asignación $\widetilde{x}$ no es socialmente eficiente. Si la economía se moviera a la asignación x', la cual es factible, ambos agentes experimentarían una mejora en su bienestar. Nótese, sin embargo, que para argumentar que $\widetilde{x}$ es ineficiente no hace falta que los dos agentes mejoren: en una asignación como x'', que también es factible, el agente 1 está estrictamente mejor y el agente 2 no ha empeorado, lo cual es suficiente para decir que $\widetilde{x}$ no era eficiente.

Caso 5 *Considere la asignación* x*:*



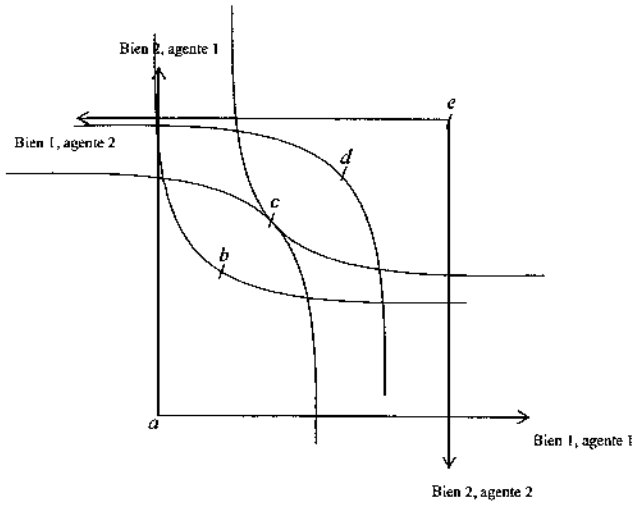
Manteniendo constante el tamaño de la caja, mejorar el bienestar del agente 1 implicaría lograr una asignación arriba/a la derecha de su curva de indiferencia, lo cual implicaría deteriorar el bienestar del agente 2. De la misma forma, mejorar la situación del agente 2 equivale a lograr una asignación abajo/a la izquierda de su curva de indiferencia, lo cual dejaría al agente 1 en una situación estrictamente peor. Esta es la situación de eficiencia social que Pareto visualizaba: es imposible lograr una mejora para algún agente de la economía sin al mismo tiempo imponer a algún otro agente un deterioro en su bienestar. Así:

Definición 2 *Una asignación (factible)* (x^1, x^2) *es eficiente en el sentido de Pareto (o es un punto de Pareto) si no existe otra asignación* $(\widehat{x}^1, \widehat{x}^2)$ *que sea factible, que dé a uno de los agentes más bienestar y dé al otro agente al menos el mismo bienestar.*

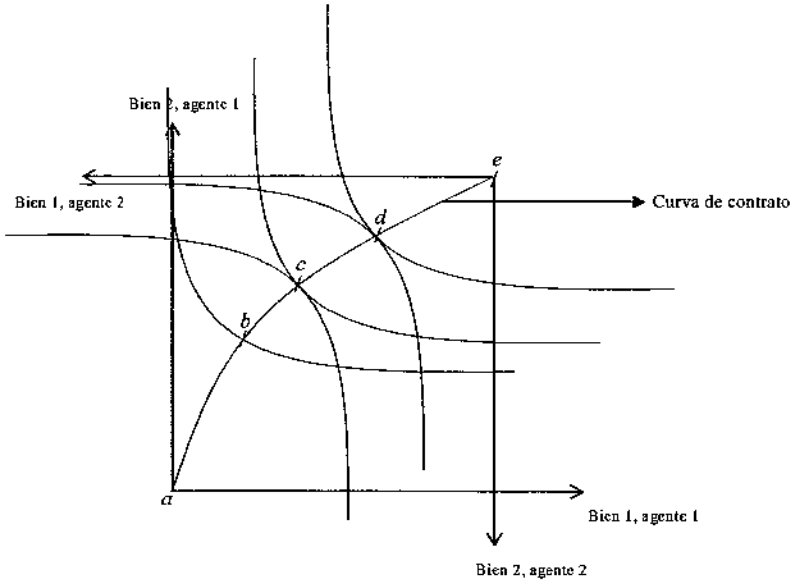
Nota técnica 3 *Nótese que para la definición de eficiencia de Pareto las dotaciones individuales no son importantes más allá de que ellas determinan el tamaño de la caja de Edgeworth (y debería ser obvio que, en general, para una caja dada hay infinidad de posibles pares de dotaciones que la generan como Caja de Edgeworth).*

1.3.2 La curva de contrato:

Es fácil ver que en una economía pueden existir muchas asignaciones diferentes que son eficientes en el sentido de Pareto. Por ejemplo: dar todo al agente 1 y nada al agente 2, o nada al 1 y todo al 2 son ambas asignaciones eficientes. Es más, con las curvas de indiferencia habituales también se suele encontrar otras masignaciones de Pareto:



En este caso, lo puntos a, b, c, d y e son todos eficientes. El conjunto de todos los puntos de Pareto de una economía es conocido como su curva de contrato:



La razón para este nombre es que es de presumir que todos los contratos de intercambio entre agentes de esta economía arrojarán asignaciones que se encuentran en esta curva; de lo contrario, al menos uno de los agentes estaría desaprovechando una oportunidad, aceptable por el otro agente, de intercambiar y lograr para sí una mejora).

Metodológicamente, es claro que en el interior de la caja de Edgeworth los puntod de Pareto son aquellos en los que las curvas de indiferencia de los agentes son tangentes. Cuando las preferencias tienen asociadas funciones de utilidad para las cuales uno puede encontrar derivadas, esta tangencia se traduce en igualdad entre las tasas marginales de sustitución:

Ejemplo 4 *Supongamos que*

$$\begin{aligned} u^1(x_1,x_2) &= \sqrt{x_1x_2} \\ u^2(x_1,x_2) &= \sqrt{x_1x_2} \\ w^1 &= (1,1) \\ w^2 &= (1,1) \end{aligned}$$

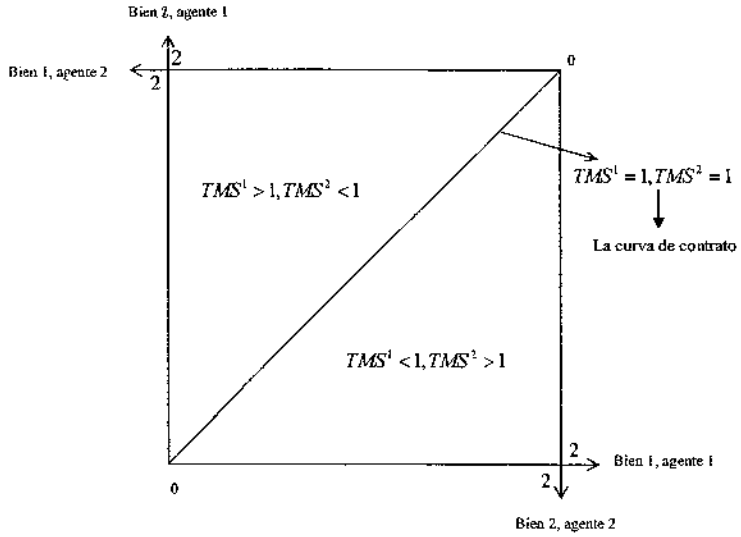
Es claro que en el borde de la Caja de Edgeworth los únicos puntos de pareto son $\left(x^1=(2,2),x^2=(0,0)\right)$ *y* $\left(x^1=(0,0),x^2=(2,2)\right)$. *Ahora, consideremos el interior de la caja. Dado que podemos diferencias las funciones de utilidad:*

$$\begin{aligned} TMS^1\left(x_1^1,x_2^1\right) &= \frac{\frac{\partial u^1}{\partial x_1}\left(x_1^1,x_2^1\right)}{\frac{\partial u^1}{\partial x_2}\left(x_1^1,x_2^1\right)} \\ &= \frac{x_2^1}{x_1^1} \end{aligned}$$

y, similarmente:

$$TMS^2\left(x_1^2,x_2^2\right)=\frac{x_2^2}{x_1^2}$$

Consideremos la caja de esta economía:



Ejercicio 3 *Encuentre la curva de contrato de las siguientes economías:*

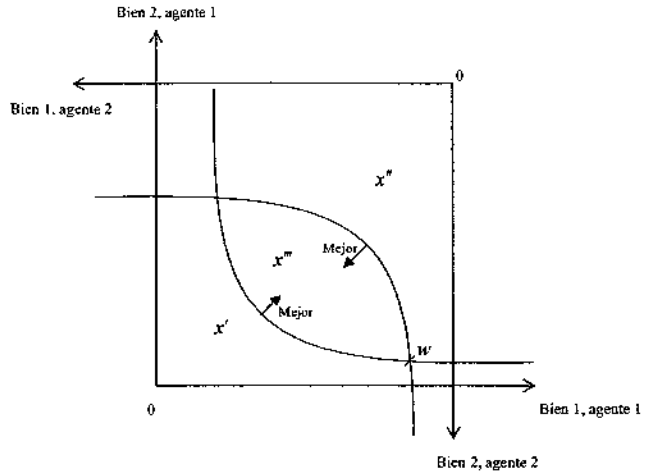
1. $u^1(x_1, x_2) = \sqrt{x_1 x_2}$, $u^2(x_1, x_2) = \sqrt{x_1 x_2}$, $w^1 = (2,1)$ *y* $w^2 = (1,1)$.
2. $u^1(x_1, x_2) = x_1^{0.6} x_2^{0.6}$, $u^2(x_1, x_2) = x_1^{0.4} x_2^{0.4}$, $w^1 = (2,2)$ *y* $w^2 = (0,0)$.
3. *(Más difícil)* $u^1(x_1, x_2) = 2x_1 + x_2$, $u^2(x_1, x_2) = x_1 + 2x_2$, $w^1 = (1,1)$ *y* $w^2 = (1,1)$

1.3.3 El análisis de Edgeworth:

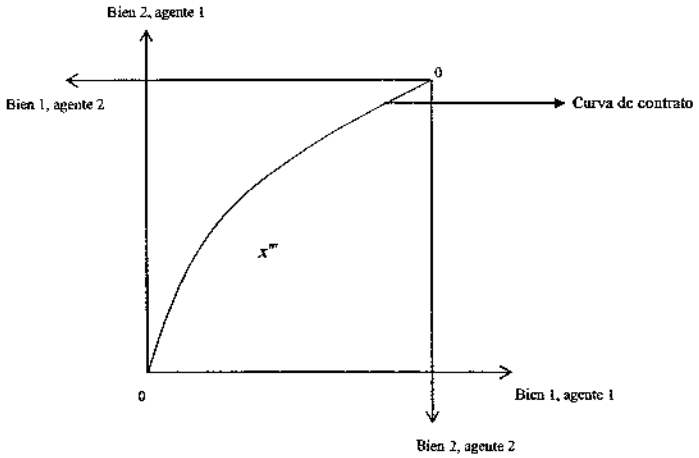
Como hemos visto con anterioridad, la definición de los puntos de Pareto de una economía sólo depende de las dotaciones de los agentes en cuanto éstas definen el tamaño de la caja de Edgeworth. Sin embargo, una vez lás dotaciones han sido establecidas, hay puntos de Pareto que uno no encontraría razonables como resultado de intercambio voluntario. Por ejemplo, si hay un agente que tiene dotación positiva de ambos bienes, no parece razonable que el intercambio lo lleve a no consumir nada de ninguno de ellos. A pesar de que la asignación en la que este individuo no consume nada es punto de Pareto, el agente no tiene por qué aceptarla: simplemente retirándose del intercambio y consumiendo su dotación, el agente lograría para sí una mejora en bienestar. Bajo esta lógica, se concluye que ningún agente debería aceptar como resultado de intercambio voluntario una asignación en la que su consumo le otorga menor bienestar que su dotación inicial.

En el siguiente gráfico, el agente 1, por simples consideraciones individuales, se opondría a una asignación como x', mientras que el agente 2, por su simple racionalidad, se opondría a una asignación como x''. Por

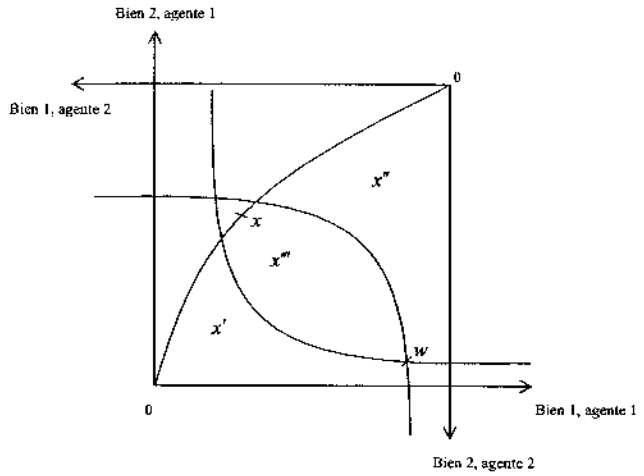
consideraciones individuales, sin embargo, ninguno de los dos se opondría a una asignación como x'''.



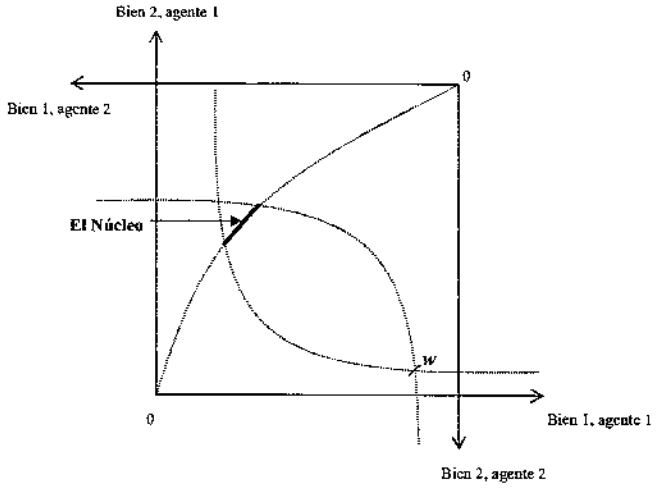
Este simple hecho no quiere decir, sin embargo, que la canasta x''' no sea objetada por nadie. Si pensamos en el par de agentes como una sociedad que se preocupa por su bienestar conjunto y, como en el próximo gráfico, resulta que x''' no está sobre la curva de contrato, sabemos que la sociedad $\{1, 2\}$ se opondría a la asignación x'''.



Ahora, consideremos la asignación x del siguiente gráfico:



Respecto a x, ninguno de los agentes en forma individual, ni la coalición $\{1,2\}$ presentarían una objeción. El conjunto de asignaciones de esta economía que satisfacen esta propiedad se denomina el Núcleo de la Economía. Gráficamente:



Formalmente:

Definición 3 *En una economía de intercambio de 2×2, el Núcleo es el conjunto de asignaciones factibles que son eficientes en el sentido de Pareto, y que cada agente encuentra superiores o indiferentes a su propia dotación.*

Ejemplo 5 *Supongamos que:*

$$\begin{aligned} u^1(x_1, x_2) &= \sqrt{x_1 x_2} \\ u^2(x_1, x_2) &= \sqrt{x_1 x_2} \\ w^1 &= (1,1) \\ w^2 &= (1,1) \end{aligned}$$

Busquemos las canastas de consumo a las que el agente 1 no tendría una objeción. Estas son, (x_1^1, x_2^1) *tal que*

$$u^1(x_1^1, x_2^1) \geq u^1(w_1^1, w_2^1)$$

es decir,

$$(x_1^1, x_2^1) : x_1^1 x_2^1 \geq 1$$

Y similarmente para el agente 2:

$$(x_1^2, x_2^2) : x_1^2 x_2^2 \geq 1$$

(evidentemente, esto descarta cualquier punto en los bordes de la Caja).

Ahora, sabemos por el ejemplo 4 que las asignaciones eficientes (a las que la coalición $\{1,2\}$ *no se opondría) satisfacen:*

$$\frac{x_2^1}{x_1^1} = \frac{x_2^2}{x_2^2}$$

mientras que por factibilidad:

$$\begin{aligned} x_1^1 + x_1^2 &= 2 \\ x_2^1 + x_2^2 &= 2 \end{aligned}$$

Esto último implica:

$$\begin{aligned} x_1^1 x_2^1 &\geq 1 \\ (2 - x_1^1)(2 - x_2^1) &\geq 1 \\ \frac{x_2^1}{x_1^1} &= \frac{2 - x_2^1}{2 - x_1^1} \end{aligned}$$

Despejando de la última de estas ecuaciones:

$$2x_2^1 - x_1^1 x_2^1 = 2x_1^1 - x_1^1 x_2^1 \Rightarrow x_1^1 = x_2^1$$

De donde se concluye que[1]

$$\begin{aligned} (x_1^1)^2 &\geq 1 \\ x_1^1 &\geq 1 \end{aligned}$$

y que[2]

$$\begin{aligned} (2 - x_1^1)^2 &\geq 1 \\ 2 - x_1^1 &\geq 1 \end{aligned}$$

Es decir:

$$1 \leq x_1^1 \leq 1$$

Por tanto, el núcleo de esta economía es el conjunto $\{x\}$ *con* $x = (x^1, x^2) = ((1,1),(1,1))$

Nota técnica 4 *Note que el núcleo depende de las dotaciones de los agentes, más allá del tamaño de la Caja de Edgeworth.*

[1] Por no negatividad, sabemos que $x_1^1 \geqslant 0$.

[2] Por factibilidad y no negatividad sabemos que $x_1^1 \leqslant 2$

Nota técnica 5 *Considere los ejemplos 3, 4 y 5. Note que, como es obvio, el punto que constituye el núcleo de la economía es un punto de Pareto. Note, además, que la asignación que resulta como equilibrio Walrasiano es también punto de Pareto. Aún más, ¡esta asignación pertenece al núcleo de la economía!.*

Ejercicio 4 *Encuentre el núcleo de las siguientes economías:*

1. $u^1(x_1,x_2)=\sqrt{x_1x_2}$, $u^2(x_1,x_2)=\sqrt{x_1x_2}$, $w^1=(2,1)$ *y* $w^2=(1,1)$.
2. $u^1(x_1,x_2)=x_1^{0.6}x_2^{0.6}$, $u^2(x_1,x_2)=x_1^{0.4}x_2^{0.4}$, $w^1=(2,2)$ *y* $w^2=(0,0)$.
3. $u^1(x_1,x_2)=2x_1+x_2$, $u^2(x_1,x_2)=x_1+2x_2$, $w^1=(1,1)$ *y* $w^2=(1,1)$

1.4 Un ejemplo:

El ejemplo que hemos venido trabajando es didáctico en el sentido de que todo es muy sencillo, pero tiene el probelma de que puede dar impresiones erróneas. A continuación trabajamos un ejemplo en el que las cosas no funcionan tan bien. Debe notarse que en este caso, aunque las preferencias no son estríctamente monótonas ni estrictamente cuasicóncavas, sí son monótonas y cuasicóncavas:

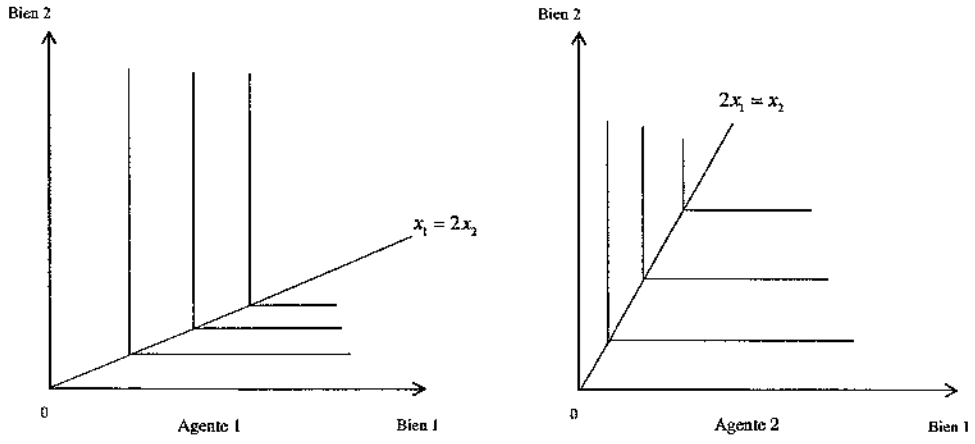
Suponemos que

$$\begin{aligned} u^1(x_1,x_2) &= \min\{x_1,2x_2\} \\ u^2(x_1,x_2) &= \min\{2x_1,x_2\} \\ w^1 &= (3,1) \\ w^2 &= (1,3) \end{aligned}$$

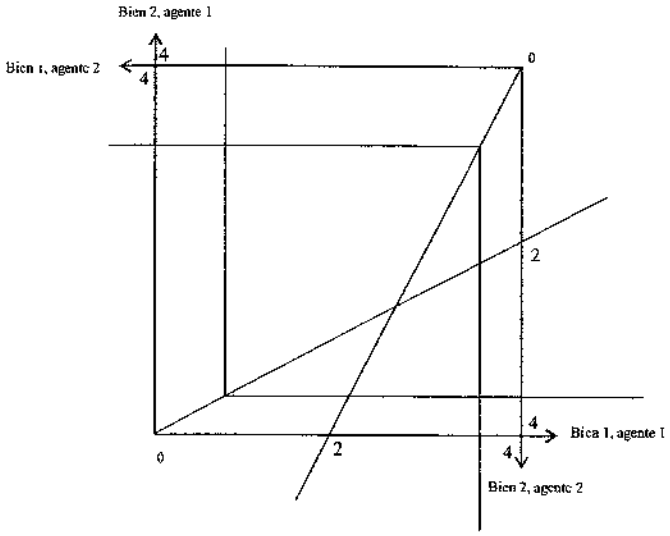
Nuestro objetivo es encontrar el equilibrio general, la curva de contrato y el núcleo, y verificar las relaciones existentes entre ellos:

1.4.1 Los mapas de indiferencia y la caja de Edgeworth:

Los mapas de indiferencia de los agentes son:



Y por lo tanto la Caja de Edgeworth es:



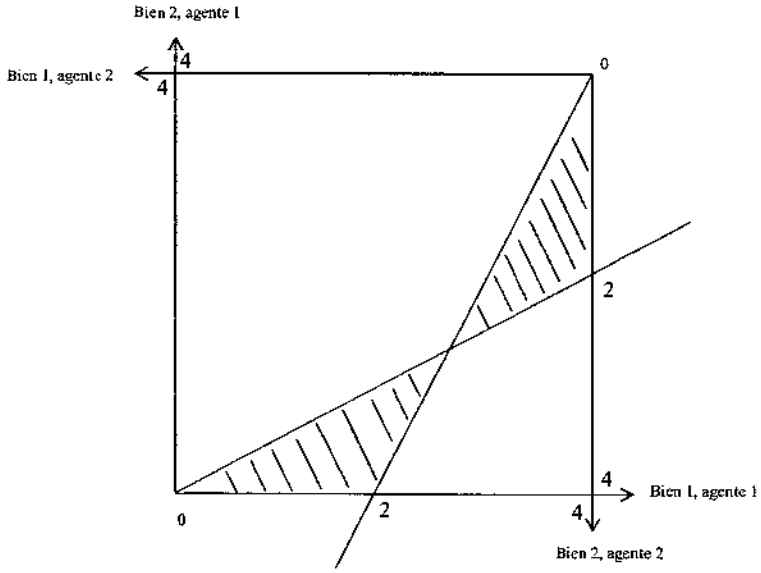
1.4.2 El equilibrio general:

Existen infinitos equilibrios generales: $\left((1,1),\left(\frac{8}{3},\frac{4}{3}\right),\left(\frac{8}{3},\frac{4}{3}\right)\right)$, $\left((1,0),\left(3,x_2^1\right),\left(1,4-x_2^1\right)\right)$ con $x_2^1 \in \left[\frac{3}{2},2\right]$ o $\left((0,1),\left(1,x_2^1\right),\left(3,4-x_2^1\right)\right)$ con $x_2^1 \in \left[2,\frac{5}{2}\right]$.

Nota técnica 6 *Nótese que hubiera sido mejor normalizar los precios a* $p_1 + p_2 = 1$.

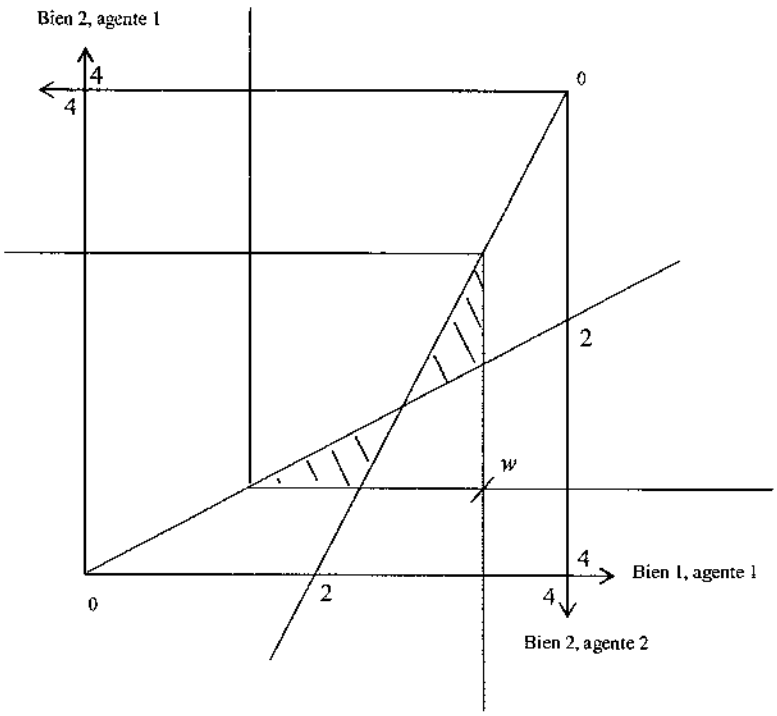
1.4.3 Curva de Contratos:

De las preferencias de los agentes es obvio que la curva de contratos es el área sombreada en el siguiente gráfico:



1.4.4 El Núcleo:

Dadas las dotaciones iniciales, se tiene que el núcleo es el área sombreada en el siguiente gráfico:



2 El problema General

Nuestro análisis gráfico de la economía 2×2 ilustra muchos conceptos, pero está muy lejos de ser realista (en particular, porque si es cierto que en una economía hay sólo dos agentes, entonces no parece muy razonable que ellos tomen los precios como dados). En esta sección vamos a generalizar nuestro modelo tanto en número de agentes como en número de bienes, y vamos a estudiar las propiedades que él implica. Por supuesto, para el modelo general la herramienta gráfica es inaplicable y tendremos que recurrir al análisis matemático. Cuando hacer esto sea razonable en términos de su complicación, lo haremos. Cuando las complicaciones del análisis formal sean demasiadas, recurriremos a la simplificación que nos brinda el análisis gráfico 2×2.

Supondremos ahora que existen $I \geq 1$ jugadores, que denotaremos por $i = 1, 2, ..., I$. Supondremos que hay $L \geq 1$ bienes, que denotaremos por $l = 1, 2, ..., L$. Como antes, supondremos que los bienes no pueden ser consumidos en cantidades negativas, de forma tal que el espacio de consumo de los agentes de $\mathbb{R}^L_+$.

Por simplicidad de notación, supondremos que cada agente i posee una función de utilidad u^i que representa sus preferencias (recordemos que lo que es verdaderamente importante es el mapa de indiferencia que estas preferencias tienen asociado, más que el valor exacto que tomen estas funciones).

Denotaremos por x^i_l el consumo del agente i del bien l. Como hemos supuesto, $x^i_l \geq 0$. Denotaremos x^i la canasta de consumo del agente i, es decir $(x^i_1, x^i_2, ..., x^i_L)$. Lo que queremos decir cuando afirmamos que u^i representa las preferencias del agente i, es que si tenemos dos canastas x y x', $u^i(x) \geq u^i(x')$ ocurre si y sólo si la canasta x es juzgada por i como indiferente o superior a la canasta x'.

Supondremos que las preferencias de los agentes satisfacen propiedades habituales:tienen funciones de utilidad contínuas, estríctamente cuasicóncavas y estríctamente monótonas.

Como nos mantendremos bajo el paradigma de la economía de intercambio, supondremos que los agentes poseen dotaciones no negativas de los bienes. Así, w_l^i será la dotación que el agente i posee del bien l. Nuestro supuesto es que $w_l^i \geq 0$. Denotaremos por w^i la canasta de bienes que el agente i posee. Como antes, $w^i = (w_1^i, w_2^i, ..., w_L^i)$.

2.1 El Análisis Walrasiano:

Denotaremos el precio del bien l por p_l. Dado el primero de los supuestos que introdujimos en la sección anterior, sólo necesitamos considerar precios estrictamente positivos, de forma que $p_l > 0$. Denotaremos por p el vector de precios $p = (p_1, p_2, ..., p_L)$.

Nuevamente, supondremos que los agentes toman los precios como dados y que la única restricción que consideran es que el valor de su consumo no supere el valor de su riqueza a los precios dados. Así, para precios p el agente i considerará como factible cualquier canasta $x = (x_1, x_2, ..., x_L)$ tal que

$$\sum_{l=1}^{L} p_l x_l \leq \sum_{l=1}^{L} p_l w_l^i$$

Una vez más, una asignación es un vector $x = (x^1, x^2, ..., x^I)$ que especifica para cada agente i una canasta de consumo x^i (en la que todos los bienes aparecen en cantidades no negativas: $x^i \in \mathbb{R}_+^L$). Una asignación es factible si no requiere más o menos que la cantidad total que hay disponible de los bienes: $x = (x^1, ..., x^I)$ es factible si

$$\sum_{i=1}^{I} x^i = \sum_{i=1}^{I} w^i$$

La intuición de lo que Walras definió como equilibrio general es la misma que antes. Equilibrio es un vector de precios y una asignación tales que, (i) para cada agente, su canasta correspondiente en la asignación es óptima desde el punto de vista de sus preferencias y su restricción presupuestal a los precios dados, y (ii) para cada bien la oferta y la demanda totales son iguales.

Aquí no podemos intentar un análisis gráfico, y debemos pasar directamente a la definición formal.

Definición 4 *Un equilibrio general para una economía como la descrita es un par compuesto por un vector de precios p y una asignación $x = (x^1, ..., x^I)$ tal que:*

(i) Cada agente maximice su utilidad a los precios dados: para todo i,

$$u^i(x^i) = \max u^i(x) \ s.a. \ \sum_{l=1}^{L} p_l x_l \leq \sum_{l=1}^{L} p_l w_l^i$$

(ii) Todos los mercados se ajusten:

$$\sum_{i=1}^{I} x^i = \sum_{l=1}^{I} w^i$$

Bajo este concepto más general, una generalización de la ley de Walras también aplica. Si la canasta x^i es óptima para el agente i a los precios p (es decir, si satisface la condición (i) de la definición anterior), entonces debe satisfacer que

$$\sum_{l=1}^{L} p_l x_l^i = \sum_{l=1}^{L} p_l w_l^i$$

Sumando para todos los agentes, obtenemos

$$\sum_{i=1}^{I}\sum_{l=1}^{L} p_l x_l^i = \sum_{i=1}^{I}\sum_{l=1}^{L} p_l w_l^i$$

lo que podemos reescribir como

$$\sum_{l=1}^{L}\left(p_l \sum_{i=1}^{I}\left(x_l^i - w_l^i\right)\right) = 0$$

Ahora, supongamos que para los $L-1$ primeros bienes el mercado se ha equilibrado, de forma tal que para todo $l = 1, 2, ..., L-1$

$$\sum_{i=1}^{I}\left(x_l^i - w_l^i\right) = 0$$

Esto implica que

$$p_L \sum_{i=1}^{I}\left(x_L^i - w_L^i\right) = 0$$

y como, bajo nuestros supuestos, $p_L > 0$, se sigue que el L-ésimo mercado tamhién debe estar en equilibrio:

$$\sum_{i=1}^{I}\left(x_L^i - w_L^i\right) = 0$$

Lo que hemos argumentado es entonces lo siguiente:

Teorema 2 *(La ley de Walras) Bajo nuestros supuestos, en una economía de intercambio con L bienes, si los agentes escogen sus demandas óptimamente, el equilibrio entre oferta y demanda en $L-1$ de los mercados implica el mismo equilibrio en el mercado restante.*

2.2 El Análisis de Pareto

El tipo de eficiencia que Pareto tenía en mente también es fácil de generalizar a cualquier número de hienes y consumidores. Recuerde que en análisis paretiano no nos ocupamos por precios, pero mantenemos que los agentes están caracterizados por preferencias y dotaciones. Con estos elementos, el concepto de eficiencia paretiana es nuevamente aquella situación en la que dada la dotación de bienes de la economía, es imposihle mejorar la situación de algún agente, sin empeorar la de algún otro.

Nuevamente, no podemos en el caso general aspirar a representar gráficamente el concepto de Punto dc Pareto o a obtener la curva de contratos. Pasamos por tanto directamente a la definición formal:

Definición 5 *En una economía como la descrita, una asignación factible $x = (x^1, x^2, ..., x^I)$ es eficiente en el sentido de Pareto (o es un punto de Pareto) si no existe otra asginación factible $\widehat{x} = (\widehat{x}^1, \widehat{x}^2, ..., \widehat{x}^I)$ tal que para todo agente i, $u^i(\widehat{x}^i) \geq u^i(x^i)$ y para al menos un agente i^*, $u^{i^*}(\widehat{x}^{i^*}) > u^{i^*}(x^{i^*})$.*

2.3 El Análisis de Edgeworth

Generalizar los conceptos de equilibrio Walrasiano y eficiencia Paretiana fué relativamente sencillo. En el caso del núcleo, las cosas ya no son tan sencillas. Ya no es suficiente que las asignaciones sean eficientes de Pareto y sean consideradas por cada agente al menos tan deseahles como su propia dotación. Recuerde que lo que Edgeworth tenía en mente con su concepto era definir aquellas situaciones ante las que ningún

agente ni grupo de agentes tendrían una objeción, en el sentido de que simplemente aislándose del intercambio podría lograr para sí mismo o para sus miembros una mejora. Cuando sólo teníamos dos agentes, definir este concepto era fácil, pues sólo teníamos que considerar a los agentes individualmente y a la sociedad como un todo. ¡Cuando hay más de dos agentes tenemos que considerar "coaliciones" que tienen más de un agente, pero que no son la sociedad completa!

Para incorporar esta dificultad, introduciremos un concepto que no habíamos definido anteriormente:

Definición 6 *La coalición de agentes* $\{i_1, i_2, ..., i_J\}$ *tiene una objeción a la asignación* $x = (x^1, x^2, ..., x^I)$ *si existe para cada uno de esos agentes una canasta* $\widehat{x}^{i_j}$ *tales que*

(i) Las canastas son factibles desde el punto de vista de la coalición:

$$\widehat{x}^{i_1} + \widehat{x}^{i_2} + ... + \widehat{x}^{i_J} = w^{i_1} + w^{i_2} + ... + w^{i_J}$$

(ii) Todos los miembros de la coalición encuentran a $\widehat{x}^{i_j}$ *al menos tan deseable como* x^{i_j}*:*

$$u^{i_1}(\widehat{x}^{i_1}) \geq u^{i_1}(x^{i_1}), u^{i_2}(\widehat{x}^{i_2}) \geq u^{i_2}(x^{i_2}), ..., u^{i_J}(\widehat{x}^{i_J}) \geq u^{i_J}(x^{i_J})$$

(iii) Al menos un miembro de la coalición i_j^* *prefiere estrictamente* $\widehat{x}^{i_j^*}$ *a* $x^{i_j^*}$*:*

$$u^{i_j^*}(\widehat{x}^{i_j^*}) > u^{i_j}(x^{i_j^*})$$

El punto del concepto de objeción es que la "coalición" de agentes se retiraría de cualquier intercambio que encuentre objetable:

(i) el agente i_j^* convencería a sus compañeros, pues él ganaría y los demás no perderían;

(ii) por otra parte, ellos pueden lograr esa mejor situación por sí mismos: no necesitan las dotaciones de los demás para lograrlo.

Con esta definición, es muy fácil definir el núcleo de una economía:

Definición 7 *El núcleo de una economía como la descrita es aquel conjunto de asignaciones factibles a las cuales ninguna coalición de agentes tiene una objeción.*

Ejercicio 5 *Demuestre que en una economía con sólo dos agentes la definición de Núcleo (definición 6) es equivalente a la definición 3.*

Ejercicio 6 *Demuestre que en cualquier economía de intercambio, el núcleo es un subconjunto del conjunto de puntos de Pareto.*

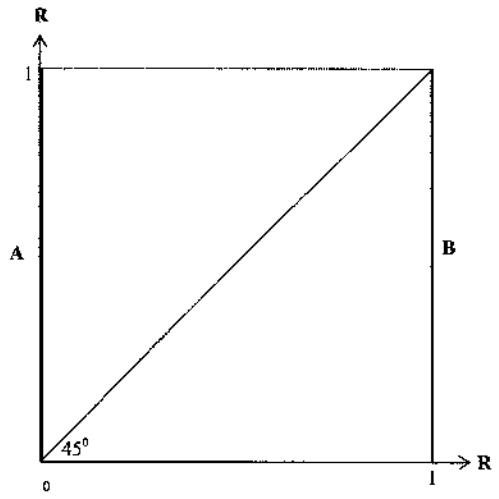
3 Análisis Positivo del Equilibrio Walrasiano:

Como hemos dicho con anterioridad, el estándar que Arrow y Debreu impusieron en economía implica que cuando quiera que una definición de equilibrio sea propuesta, esta se acompañe de un profundo estudio de sus propiedades positivas. En un mundo ideal, nuestro equilibrio habría de satisfacer lo siguiente: (i) siempre existir; (ii) ser único; (iii) ser siempre estable bajo algún mecanismo natural de ajuste; (iv) permitirle a uno refutar la bipótesis de equilibrio con base en evidencia empírica. En esta sección nos vamos a dedicar a estudiar estos aspectos, en la medida en la que la complejidad técnica nos lo permita.

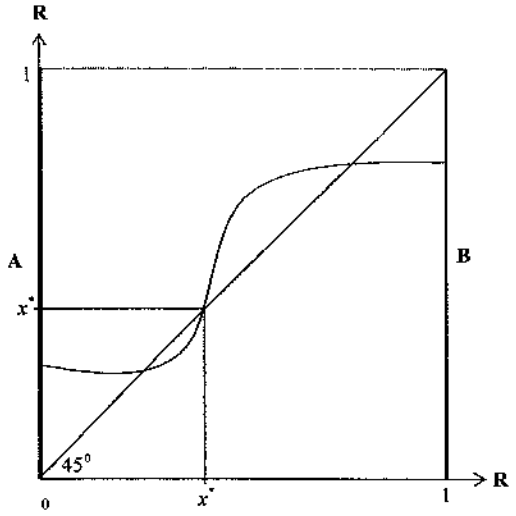
3.1 Existencia:

3.1.1 Una introducción a los Teoremas de Punto Fijo:

Considere el siguiente problema: dado el siguiente gráfico,



intentemos trazar una gráfica contínua que conecte el tramo A con el tramo B y no cruce la diagonal de 45°. Esto es evidentemente imposible. Como en el siguiente gráfico,



cualquier gráfica contínua tendrá por lo menos un punto en el cual la coordenada horizontal será igual a la coordenada vertical. Esto es:

Teorema 3 *Para cualquier función* $f : [0,1] \to [0,1]$ *que sea contínua, existe* $x^* \in [0,1]$ *tal que* $f(x^*) = x^*$.

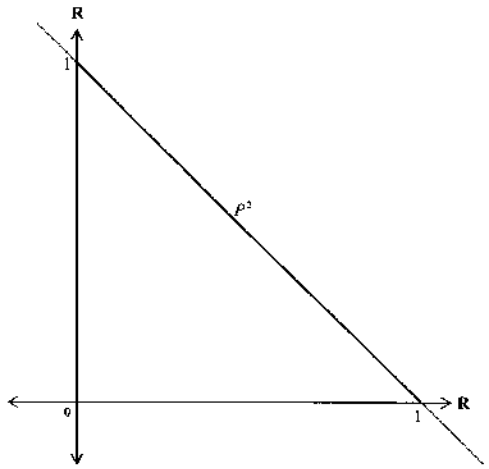
Este ejemplo ha sido didáctico, pero no tiene la generalidad que uno quisiera. En general, f no tiene que ser una función que relacione $[0,1]$ con sí mismo. Lo que necesitamos requerir es lo siguiente: (i) que

la función sea contínua; (ii) que el dominio y el codominio sean el mismo conjunto; (iii) que este conjunto tenga un principio y un fin, es decir, que no vaya hasta infinito o hasta menos infinito (acotado); (iv) que el conjunto contenga sus puntos límite, o su borde (cerrado); (v) que uno pueda trazar una línea entre dos puntos cualesquiera de este conjunto, sin que ésta se salga del conjunto (convexo). Con estas propiedades, uno siempre encontrará un punto en el dominio que sea igual a su imágen bajo f.

Por ejemplo, si tomamos el siguiente conjunto:

$$P^2 = \{(p_1, p_2) \in \mathbb{R}^2_+ \mid p_1 + p_2 = 1\}$$

que tiene la siguiente forma



y es por tanto, acotado, cerrado y convexo, y trazamos una función cualquiera,

$$f : P^2 \to P^2$$

que sea contínua, siempre existirá (p_1^*, p_2^*) tal que

$$f(p_1^*, p_2^*) = (p_1^*, p_2^*)$$

Esto lo podemos hacer para cualquier dimensión: si tenemos $L \geq 1$ bienes, y definimos

$$P^L = \{(p_1, p_2, ..., p_L) \in \mathbb{R}^L_+ \mid \sum_{l=1}^{L} p_l = 1\}$$

entonces,

Teorema 4 *Para cualquier función $f : P^L \to P^L$ que sea contínua, existe un punto $p^* = (p_1^*, p_2^*, ..., p_L^*)$ tal que*

$$f(p^*) = p^*$$

3.1.2 El Subastador Walrasiano:

Walras tenía en mente un proceso de ajuste de precios que correspondería a lo intuitivamente obvio: si un mercado muestra exceso de demanda, su precio (relativo) debe subir, y si un mercado muestra exceso de oferta, su precio debe bajar.

Nótese, sin embargo, un gran problema que tiene el modelo de equilibrio general: en física, uno comienza el análisis son un proceso dinámico (por ejemplo, la moción de un péndulo) y el equilibrio aparece, de manera natural, como la situación en la que las fuerzas dinámicas se cancelan. En equilibrio general, como en muchas otras teorías económicas, ¡formulamos conceptos de equilibrio que no provienen de una dinámica previa!. Puesto en otras palabras: Equilibrio General nos describe precios ante los cuales no pareceríamos tener por qué movernos, pero no nos dice cómo llegamos a ellos.

Lo más cercano que tenemos a un proceso dinámico de ajuste fué un artefacto que Walras denominó el Tatonador (subastador) y que hacía exactamente lo que hemos dicho previamente: mover los precios en la dirección indicada por los excesos de demanda.

Así, supongamos una ecouomía como la descrita en la sección anterior. Denotemos, para precios normalizados $p \in P^L$, la demanda del agente i por $x^i(p)$. La demanda agregada a precios p se denotará por $x(p)$:

$$x^A(p) = \sum_{i=1}^{I} x^i(p)$$

y la función de exceso de demanda a precios p:

$$z(p) = (z_1(p), z_2(p), ..., z_L(p)) = x^A(p) - \sum_{i=1}^{I} w^i$$

Lo que el subastador va a hacer es suhir los precios de aquellos bienes l para los cuales $z_l(p) > 0$ y bajar los precios de aquellos para los cuales $z_l(p) < 0$ (exceso de oferta).

La forma más sencilla de lograr esto sería la siguiente: ante los precios p el subastador reaccionaría definiendo los nuevos precios:

$$p' = p + z(p)$$

Aquí, sin embargo, el subastador encuentra un problema, para un bien con un "gran" exceso de oferta, el precio p'_l que él definiría sería negativo. Esto no puede ser, y por lo tanto él decide lo siguiente: si el precio p'_l que resulta con la regla anterior es no negativo, entonces él deja ese p'_l. Si llega a ser negativo, entonces el nuevo precio será cero. Esto es, ante precios p, él reacciona con los siguientes precios: para el bien l,

$$p'_l = Max\{p_l + z_l(p), 0\}$$

Abora, el subastador enfrenta un nuevo problema: estos precios p no necesariamente estarán normalizados (como sí lo estará por supuesto p). Esto lo soluciona el subastador de manera muy fácil. Definamos para cada l una función:

$$\varphi_l(p) = Max\{p_l + z_l(p), 0\}$$

entonces, ante precios p el subastador reacciona con

$$p'_l = \frac{\varphi_l(p)}{\sum_{l'=1}^{L} \varphi_{l'}(p)}$$

Nota técnica 7 *Para poder hacer esto, necesitamos tener garantía de que* $\sum_{l=1}^{L} \varphi_l(p) > 0$. *Debería ser claro que la ley de Walras tiene esta implicación.*

Por construcción, si $p' = (p'_1, p'_2, ..., p'_L)$ es obtenido a partir de la anterior ecuación, tenemos que $p' \in P^L$. Por tanto, podemos hacer explícita la notación, definiendo una función $f : P^L \to P^L$, donde,

$$\begin{aligned} f(p) &= (f_1(p), f_2(p), ..., f_L(p)) \\ &= (\frac{\varphi_1(p)}{\sum_{l=1}^{L} \varphi_l(p)}, \frac{\varphi_2(p)}{\sum_{l=1}^{L} \varphi_l(p)}, ..., \frac{\varphi_L(p)}{\sum_{l=1}^{L} \varphi_l(p)}) \end{aligned}$$

3.1.3 Existencia del Equilibrio General:

Utilizando la definición anterior del subastador Walrasiano, es relativamente sencillo demostrar que, bajo los supuestos habituales, el equilibrio Walrasiano siempre existe.

El punto es el siguiente: bajo los supuestos habituales, cada función de demanda x^i es contínua (es suficiente que las preferencias de i tengan asociada una función de utilidad contínua). Esto implica que x^A, y por tanto z, son funciones contínuas, lo cual implica que la función f del subastador es también contínua.

Por el teorema 4, entonces sabemos que existe $p \in P^L$ tal que

$$p = f(p)$$

Ignoremos por simplicidad el denominador de la definición de f, lo cual no es un problema, por la homogeneidad de las demandas. Lo que hemos encontrado es que para cada l,

$$p_l = \max\{p_l + z_l(p), 0\}$$

Si $p_l > 0$, tiene que ser porque

$$p_l = p_l + z_l(p)$$

lo cual implica que $z_l(p) = 0$.

Si, por otra parte, $p_l = 0$, entonces,

$$p_l = 0 \text{ y } z_l(p) \leq 0$$

Bajo nuestros supuestos esto no tiene lógica: un bien que es gratis estaría en exceso de oferta. Descartando esta posibilidad, hemos encontrado que a los precios de p, $z(p) = 0$. ¡Estos son precios de equilibrio general!

Nota técnica 8 *Nuestra prueba tiene un pequeño problema cuando alguno de los precios se hace cero (la demanda puede no estar definida). Uno puede evitar este problema, pero esto es un argumento fundamentalmente técnico que no hace falta estudiar aquí. Además, aquí hemos implícitamente que las demandas de los agentes a precios dados están definida de manera única. Con teoremas de punto fijo más potentes se puede tratar el problema cuando las preferencias simplemente son cuasicónavas, dado que la demanda en este caso será de valores convexos y hemicontínua superior..*

3.2 Otros aspectos positivos del modelo de Equilibrio General:

Como acabamos de ver, el resultado positivo más importante que puede demandarse de un concepto de equilibrio, sin existencia, es obtenido bajo los supuestos habituales. La siguiente pregunta que uno suele formularse en términos de análisis positivo es la de unicidad, y luego algo que suele ser relacionado, la de estabilidad. Infortunadamente, las técnicas que necesitaríamos utilizar para el tratamiento profundo de estos temas están muy por encima del nivel de este curso, y nos contentaremos con un recuento simple de estos resultados.

Antes de analizar estos aspectos positivos, introduciremos el siguiente resultado:

3.2.1 El Teorema SMD:

De manera independiente, tres economistas muy famosos, Hugo Sonnenschein, Rolf Mantel y el ya conocido Gerard Debreu, estudiaron las propiedades que la estructura habitual del modelo de equilibrio general impone en la función de exceso de demanda agregada, Z, que introdujimos previamente.

El resultado fué negativo, en el sentido de qye prácticamente ninguna propiedad es impuesta.

Supongamos que definimos el conjunto de precios normalizados P como antes. Debería ser obvio que para todo $p \in P$,

$$\sum_{l=1}^{L} p_l Z_l(p) = 0$$

(es decir, la Ley de Walras).

El resultado que Sonnenschein, Mantel yDebreu obtuvieron puede resumirse como que esa es la única propiedad que una función de exceso de demanda tiene que satisfacer.

En particular, ellos demostraron que si a uno le dan una función $z : P \to \mathbb{R}^n$ tal que satisfaga la ley de Walras: para todo $p \in P$,

$$\sum_{l=1}^{L} p_l z_l(p) = 0$$

Uno siempre puede encontrar una economía (es decir, unos consumidores descritos por unas preferencias y unas dotaciones) tal que z es la función de exceso de demanda agregada de esa economía (esto ocurre en todo P salvo cuando uno está muy cerca de precios que se hacen cero). La única condición es que uno pueda construir al menos tantos consumidores como bienes.

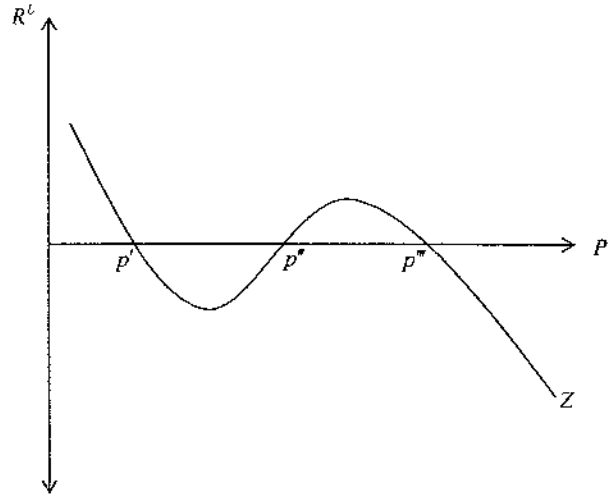
Este resultado se conoce como el teorema SMD (o SDM). A continuación vamos a utilizar este resultado para estudiar las otras propiedades positivas del modelo de Equilibrio General.

3.2.2 Unicidad:

Lo primero que se concluye del SMD es que no hay por qué esperar que el Equilibrio General sea único: no hay ninguna razón por la cual la función Z de una economía sólo deba tener un punto p en P en el cual $Z(p) = 0$.

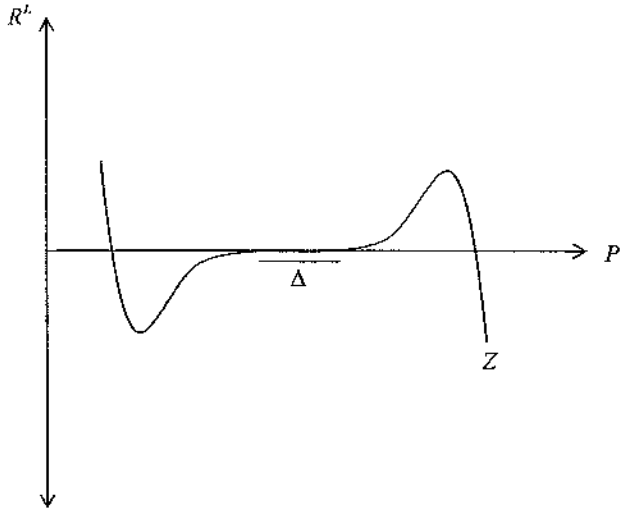
Por ejemplo, ignorando la dimensionalidad del problema, uno podría tener que Z luce como en el siguiente

gráfico:



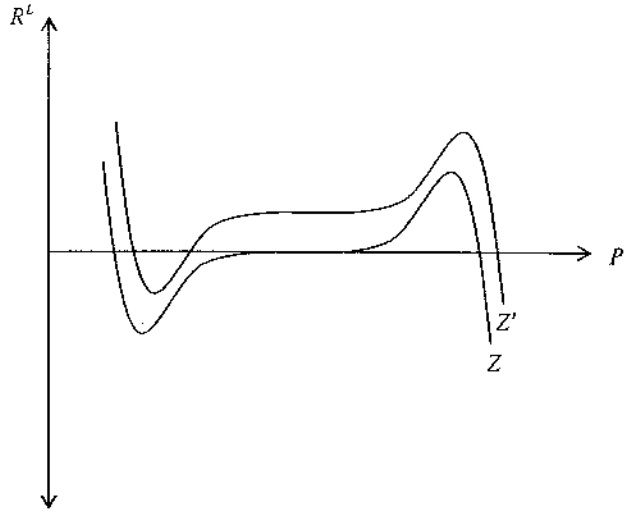
En cuyo caso tendríamos tres precios (normalizados) de equilibrio. Esto no debería resultar sorprendente: por una parte, nosotros ya hemos obtenido multiplicidad de equilibrios en algunos de nuestros ejemplos; por otra, los teoremas de punto fijo, como el que utilizamos para demostrar existencia, aseguran que un punto fijo existe pero no que es único.

Como hemos dicho con anterioridad, Arrow demostró que las condiciones para que el equilibrio sea único son extremas (básicamente, que todos los agentes sean Cobb-Douglas). Debreu, sin embargo, estaba preocupado por un problema más complicado: nada garantiza que la función Z no luzca como a continuación:



En cuyo caso uno tendría un número infinito de equilibrios, y lo que es más grave, no encontraría todo un "contínuo" de equilibrios, como Δ. El problema sería que en una situación como esta los equilibrios ni siquiera son únicos en un sentido local: ¡se encuentran infinitamente cerca! Debreu demostró, sin embargo, que esto "casi nunca" pasa. Lo siguiente sobre-simplifica el análisis, pero supongamos que las dotaciones de la economía no hubieran sido las que generaron el gráfico Z anterior, sino otras levemente diferentes. Uno

entonces, esperaría una función Z' como a continuación:

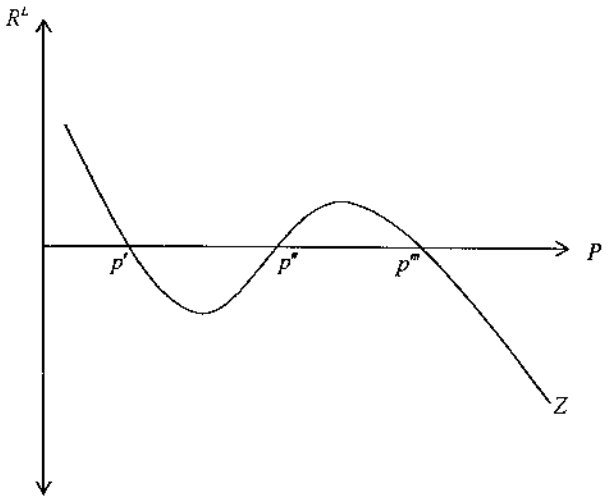


donde sólo un número finito de equilibrios (todos ellos aislados) se presenta. Necesitaría uno una tremenda coincidencia para que fuera Z y no Z' la función de demanda agregada de la economía.

En síntesis, no hay razón para esperar que el Equilibrio General sea único, pero casi siempre uno encuentra que es localmente único.

3.2.3 Estabilidad:

Como ya hemos dicho, la definición de Equilibrio General carece de un mecanismo natural que explique cómo la economía se desenvuelve cuando uno se encuentra por fuera de equilibrio. Hemos propuesto un mecanismo de Subastador Walrasiano, que, sin ser natural, parece aceptable. Nótese, sin embargo, que bajo este mecanismo el equilibrio general no tiene por qué ser estable (aún localmente). Como vimos anteriormente, del SMD se sigue que Z puede ser como a continuación:



P'', aún siendo equilibrio general, ¡no es estable bajo el subastador Walrasiano!

3.2.4 Refutabilidad:

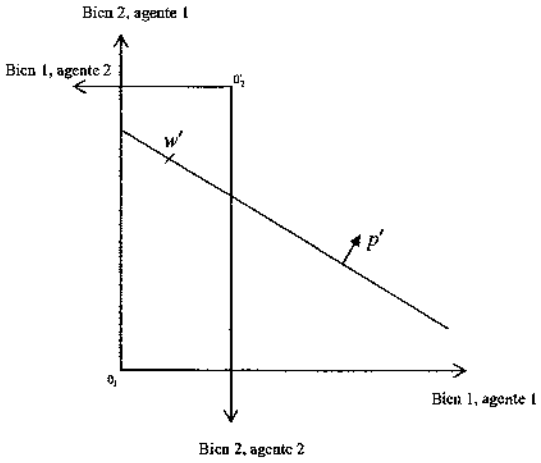
Note cómo hemos utilizado hasta ahora el SMD: hemos aprovechado el resultado para argumentar que no podemos descartar funciones de exceso de demanda agregada, a pesar de lo "mal comportadas" que éstas puedan resultar. Pareciera como si cualquier cosa fuera compatible con el equilibrio general, como si uno nunca pudiera refutar la hipótesis de equilibrio general.

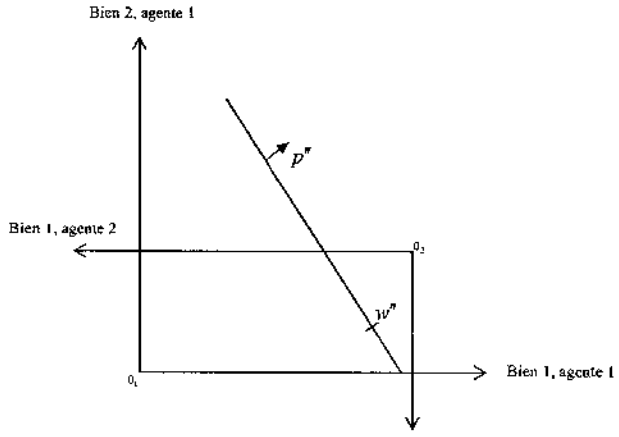
Esto resultaría problemático, pues según una importante corriente epistemológica conocida como "falsificacionismo", sólo las ideas que son refutables son conocimiento científico.

En efecto, durante mucho tiempo los economistas creíamos que del SMD se desprendía que la hipótesis de equilibrio general no era refutable.

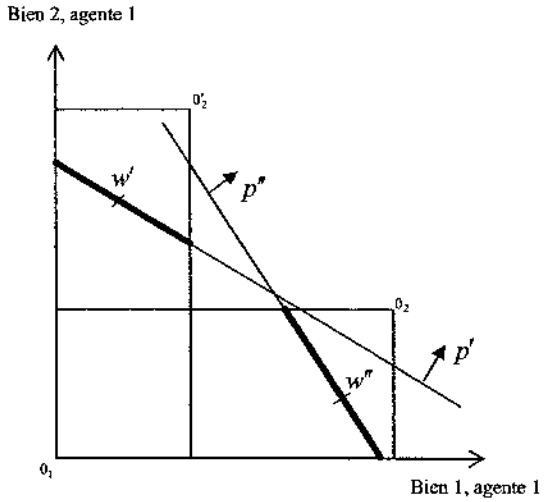
Muy recientemente se ha demostrado que esto no es así. Cuando uno utiliza el SMD está manteniendo las dotaciones fijas y permitiendo sólo a los precios variar. Donald Brown y Rosa Matzkin han demostrado que si uno permite que las dotaciones varíen, uno puede refutar la hipótesis de EG.

Por ejemplo, tomemos una economía 2×2 en la que se han observado las siguientes dotaciones y precios:





Sobreponiendo los dos gráficos, obtenemos las asignaciones que serían factibles en cada caso como equilibrio general (las partes más gruesas de cada gráfico).



Claramente, no es posible que el agente 1 satisfaga el axioma débil de las preferencias revcladas. Esto implica que es imposible que dadas las observaciones de dotaciones y precios, al mismo tiempo los agentes maximicen su bienestar y los mercados se vacíen: ¡es posible refutar la hipótesis de Equilibrio General!.

4 Análisis Normativo del Equilibrio Walrasiano:

4.1 Los Teoremas Fundamentales de Economía del Bienestar:

Hemos visto que un concepto de bienstar social relativamente poco controversial desde el punto de vista ético es el de eficiencia de Pareto: lo mínimo que una asignación tendría que satisfacer para ser caeptable por una sociedad es ser eficiente de Pareto. En esta sección analizamos las relaciones existentes entre las asignaciones de equilibrio Walrasiano y la eficiencia de Pareto.

4.1.1 El primer teorema:

Recordemos los ejemplos 3 y 4 y los ejercicios que les siguieron. Allí observamos que las asignaciones de equilibrio competitivo eran todas eficientes de Pareto. Nuestro primer resultado es que esto no es una coincidencia:

Teorema 5 El primer teorema fundamental de economía del bienestar: *Dada una economía como la que hemos descrito, si* $\left(p,\left(x^1,x^2,...,x^I\right)\right)$ *es un equilibrio general, entonces* $\left(x^1,x^2,...,x^I\right)$ *es eficiente en el sentido de Pareto.*

Muchos de los resultados de la teoría del equilibrio general requieren métodos matemáticos muy complejos para su demostración formal. Este teorema, a pesar de su importancia, es una notable excepción, razón por la cual a continuación presentamos su prueba completa. El método de prueba que vamos a utilizar se conoce como "prueba por contradicción."

Prueba. Supongamos que $\left(p,\left(x^1,x^2,...,x^I\right)\right)$ es un equilibrio general, pero $\left(x^1,x^2,...,x^I\right)$ **no** es eficiente en el sentido de Pareto. Entonces, existe $\left(\widehat{x}^1,\widehat{x}^2,...,\widehat{x}^I\right)$ tal que:

1. $\sum_{i=1}^{I}\widehat{x}^i=\sum_{i=1}^{I}w^i$

2. Para todo i, $u^i\left(\widehat{x}^i\right) \geqslant u^i\left(x^i\right)$

3. Para algún i^*, $u^{i^*}\left(\widehat{x}^{i^*}\right) > u^{i^*}\left(x^{i^*}\right)$

Por definición de equilibrio, se sigue de la condición 3 que $p \cdot \widehat{x}^{i^*} > p \cdot x^{i^*}$, mientras que la 2 implica que para todo i, $p \cdot \widehat{x}^i \geqslant p \cdot x^i$. Sumando para todos los agentes, obtenemos que

$$\sum_{i=1}^{I} p \cdot \widehat{x}^i > \sum_{i=1}^{I} p \cdot x^i$$

de donde se deduce que

$$p \cdot \sum_{i=1}^{I} \widehat{x}^i > p \cdot \sum_{i=1}^{I} x^i = p \cdot \sum_{i=1}^{I} w^i$$

lo cual contradice la condición 1. ■

El punto de la demostración es simple: asignaciones que serían preferibles para los consumidores deben costar más y, por tanto no pueden ser factibles si nos encontramos en equilibrio general.

El teorema es de fundamental importancia para las autoridades de política económica: si una economía satisface los supuestos del modelo de equilibrio Walrasiano y se encuentra en equilibrio, las medidas de política económica que pretendan mejorar el bienestar de algún individuo, manteniendo las dotaciones fijas, necesariamente irán en detrimento del bienestar de alguien más. Esta idea no es para nada nueva: ¡es lo que Adam Smith llamaba la "Mano Invisible"!

Nota técnica 9 *En la demostración hemos utilizado el supuesto de que cualquier cantidad positiva, por pequeña que sea, de cualquier bien mejora el bienestar de cualquier agente. Esto lo hemos hecho para descontar la posibilidad de que los agentes tengan curvas de indiferencia gruesas. Un ejercicio interesante sería mostrar que, aún en una economía* 2×2*, si un agente tiene curvas de indiferencia gruesas la conclusión del teorema no se cumple.*

4.1.2 El segundo teorema:

En el primer teorema la pregunta es si una asignación de equilibrio es eficiente. El segundo teorema se plantea de alguna manera la pregunta inversa: si tomamos una asignación eficiente, ¿podemos garantizar que esta sea de equilibrio competitivo?

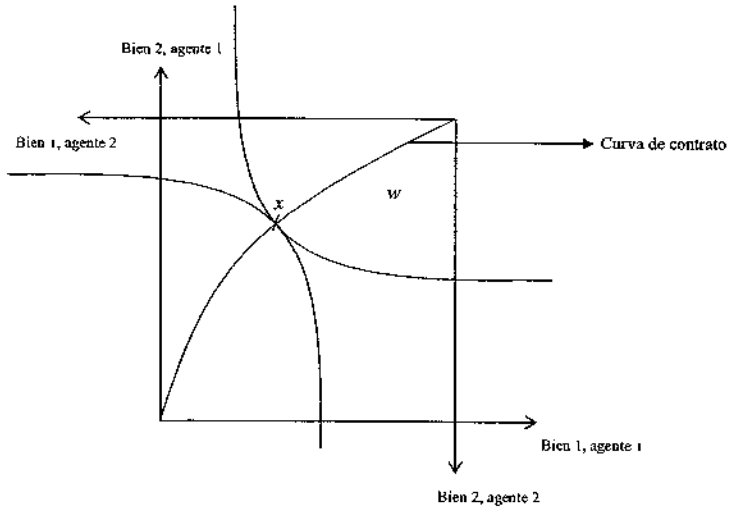
La respuesta a la pregunta así planteada es obviamente negativa: dadas unas dotaciones, ya sabemos que hay asignaciones que, aun siendo eficientes, no podrían resultar de intercambios voluntarios. El punto está en si mantenemos las dotaciones iniciales fijas o permitimos redistribuciones de ellas (que no alteren la dotación agregada). El segundo teorema fundamental de economía del bienestar resuelve este problema: bajo nuestros supuestos, si permitimos redistribución de las dotaciones iniciales, entonces cualquier asignación eficiente puede ser sostenida en un equilibrio general. Formalmente:

Teorema 6 El segundo teorema fundamental de economía del bienestar: *Dada una economía constituída por preferencias* $\left(u^1, u^2, ..., u^I\right)$ *y dotaciones* $\left(w^1, w^2, ..., w^I\right)$*, si* $\left(x^1, x^2, ..., x^I\right)$ *es una asignación eficiente entonces existe una redistribución de las dotaciones* $\left(\widehat{w}^1, \widehat{w}^2, ..., \widehat{w}^I\right)$ *y unos precios* $p = (p_1, p_2, ..., p_L)$ *tales que:*

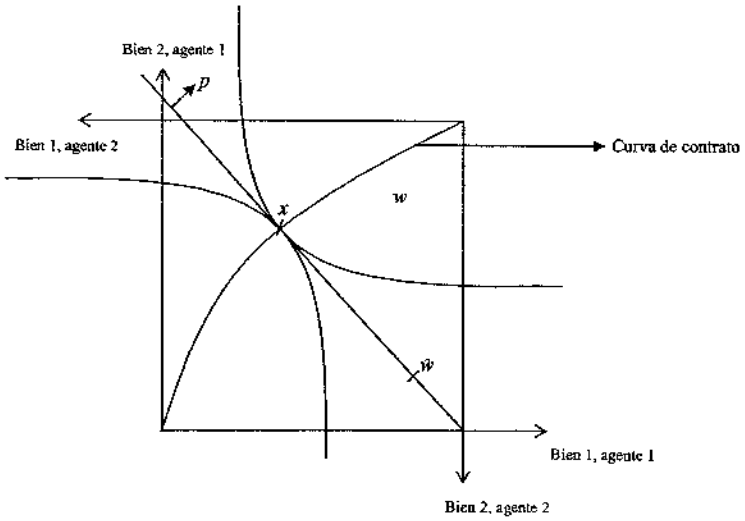
1. $\sum_{i=1}^{I} \widehat{w}^i = \sum_{i=1}^{I} w^i$
2. $\left(p, \left(x^1, x^2, ..., x^I\right)\right)$ *es un equilibrio Walrasiano de la economía constituída por preferencias* $\left(u^1, u^2, ..., u^I\right)$ *y dotaciones* $\left(\widehat{w}^1, \widehat{w}^2, ..., \widehat{w}^I\right)$.

La implicación del teorema es que si una autoridad de política económica desea imponer una asignación eficiente, no necesita cerrar los mercados. Por el contrario, puede limitarse a redistribuir las dotaciones (política fiscal) de manera adecuada y luego permitirle a los mercados actuar, pues éstos deberían llevar a la economía a la asignación descada.

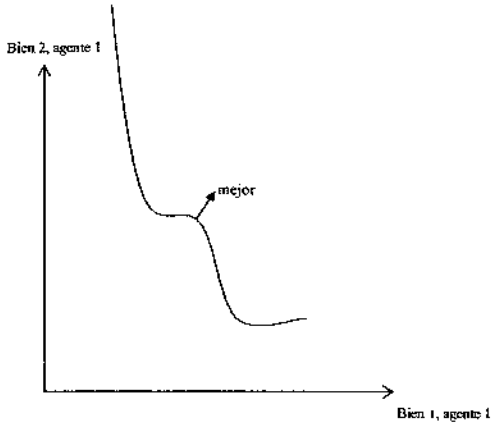
Infortunadamente, y en contraste con el caso de primer teorema, la prueba de este segundo teorema es muy compbcada y nos vamos a limitar a una ilustración gráfica en el caso 2×2. El argumento es simple. Considere la siguiente gráfica:



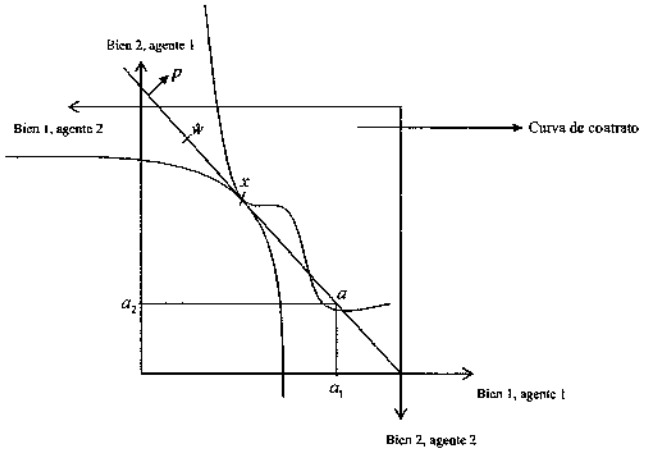
La asignación x, aunque eficiente, no puede ser de equilibrio con las dotaciones iniciales w (¿por qué no?). Sin embargo, si trazamos la tangente a ambas curvas de indiferencia en el punto x, es claro que simplemente con redistribuir las dotaciones a un punto sobre esta recta, como por ejemplo $\widehat{w}$, o el mismo punto x, obtenemos que los precios p, que están implícitos en la pendiente de esta tangente, y la asignación x son equilibrio bajo las nuevas dotaciones.



Este teorema tampoco está libre de supuestos. En este caso, la forma de las curvas de indiferencia es clave. Si, contrario a nuestros supuestos, las curvas de indiferencia del agente 1 fueran como a continuación,



entonces la conclusión del teorema no aplicaría, como podemos ver en el siguiente gráfico:



Aquí, la asignación $x = (x^1, x^2)$ es punto de Pareto. Para "convencer" al agente 2 de demandar x^2, necesitaríamos unas dotaciones como $\widehat{w}$ y los precios p. Sin embargo, el agente 1 no querría demandar x^1 bajo su restricción presupuestal correspondiente, pues una canasta como $a = (a_1, a_2)$ le resultaría superior y factible.

4.2 El Equilibrio General y el Núcleo:

4.2.1 Pertenencia:

Otra observación casual de los ejemplos 3 y 5 (y los ejercicios que les siguieron) fue que, aparentemente, las asignaciones de equilibrio pertenecen al núcleo de la economía. Esto es efectivamente así:

Teorema 7 *Dada una economía como la que hemos descrito, si* $(p, (x^1, x^2, ..., x^I))$ *es un equilibrio general, entonces* $(x^1, x^2, ..., x^I)$ *pertenece al núcleo de la economía.*

Prueba. Si uno entiende la prueba del teorema 5, le debe resultar sencillo demostrar este resultado. Esta prueba queda, por tanto, como ejercicio. ■

Teorema 8 *Dada una economía de intercambio constituída por preferencias* $\left(u^1, u^2, ..., u^I\right)$ *y dotaciones* $\left(w^1, w^2, ..., w^I\right)$, *si la asignación* $\left(x^1, x^2, ..., x^I\right)$ *está en el núcleo de la economía, entonces existe una redistribución de las dotaciones* $\left(\widehat{w}^1, \widehat{w}^2, ..., \widehat{w}^I\right)$ *y unos precios* $p = (p_1, p_2, ..., p_L)$ *tales que:*

1. $\sum_{i=1}^{I} \widehat{w}^i = \sum_{i=1}^{I} w^i$
2. $\left(p, \left(x^1, x^2, ..., x^I\right)\right)$ es un equilibrio Walrasiano de la economía constituída por preferencias $\left(u^1, u^2, ..., u^I\right)$ y dotaciones $\left(\widehat{w}^1, \widehat{w}^2, ..., \widehat{w}^I\right)$.

Prueba. Queda como ejercicio. ■

Lo que el teorema quuiere decir es que ningún individuo ni grupo de individuos puede obtener para sí o para sus miembros una mejora, simplemente con aislarse del intercambio de equilibrio. Puesto de otra forma, el resultado es que si, estando en una situación de equilibrio, uno cerrara los mercados y les propusiera a los agentes que hicieran trueques para mejorar sus situaciones, no habría ningún incentivo para estos trueques: en el equilibrio, las posibles ganancias del intercambio se han agotado.

4.2.2 Convergencia:

Hemos encontrado que las asignaciones de equilibrio pertenecen al núcleo, pero también habíamos visto en nuestros ejemplos que hay asignaciones en el nucleo que no son de equilibrio competitivo para unas asignaciones dadas.

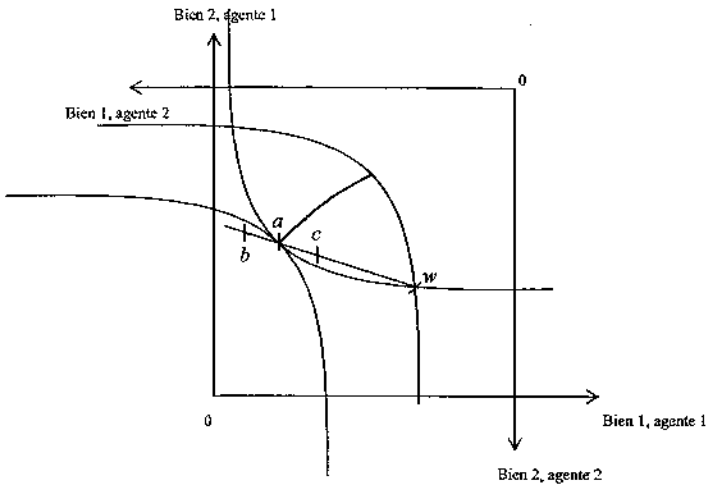
Lo que vamos a intentar ahora, sin embargo, va en otra dirección. Nótese que de la misma definición de núcleo uno intuye que al aumentar el número de agentes de la economía el núcleo de ésta se "reduce," pues al entrar nuevos agentes hay más coaliciones que pueden presentar objeciones.[3] Debreu y Herbert Scarf se plantearon este preciso problema: (i) ¿es cierto que el núcleo se "reduce" cuando aumenta el número de agentes? y (ii) si llevamos a infinito el número de agentes, ¿quedarán en el núcleo asignaciones que no son de equilibrio?

Para una forma muy particular de aumentar el número de agentes, las respuestas a las anteriores preguntas fueron, respectivamente, afirmativa y negativa.

El aumento de agentes que ellos consideraron fue simplemente la replicación de los agentes existentes. Su resultado es muy difícil de explicar formalmente, pero a continuación voy a dar un argumento intuitivo que

[3] Utilizamos la palabra "reduce" en un sentido intuitivo, pues al entrar más agentes cambia la dimensión del espacio en el cual está definido el núcleo.

está muy lejos de ser perfecto. Consideremos una economía de 2×2, como la hahitual:



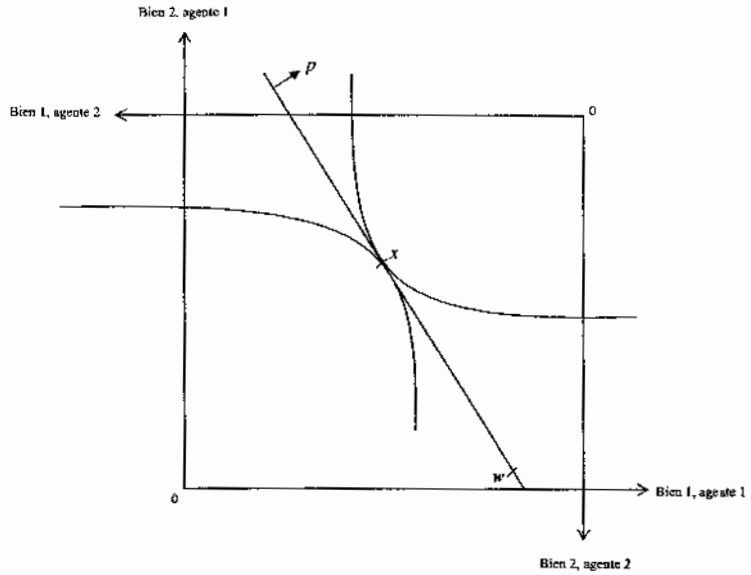
El punto a está en el núcleo. Sin embargo, los agentes preferiríau una situación en la que el agente 1 consume en c y el agente 2 en b. ¿Por qué no tienen ellos estos consumos? Porque no es posible: habría exceso de demanda por el bien 1 y exceso de oferta del 2.

Ahora, si a esta economía llegara una réplica del agente 1, quien tiene "mucho" del bien 1 y "poco" del bien 2, en el sentido de que cuando él consume en c es oferente neto de 1 y demandante neto de 2, entonces una asignación como (c, c, b) sería factible para la coalición $\{1, \text{réplica de } 1, 2\}$ y sería más deseable que la que está implícita en el punto a. La existencia de la réplica del agente 1 "eliminaría" la asignación a del núcleo.

4.3 La paradoja de las transferencias:

Finalmente, nos ocupamos de un resultado que muestra cómo las políticas económicas pueden arrojar resultados diferentes a su propósito, si no se acompañan de un buen entendimiento de la economía. Este tipo de resultados fue inicialmente propuesto por Edgeworth, pero tuvo muy poca aceptación, pues iba en contra de la intuición común de la época. Hoy sabemos que las ideas de Edgeworth eran correctas.

Consideremos una situación de equilibrio en una economía 2×2 como a continuación.



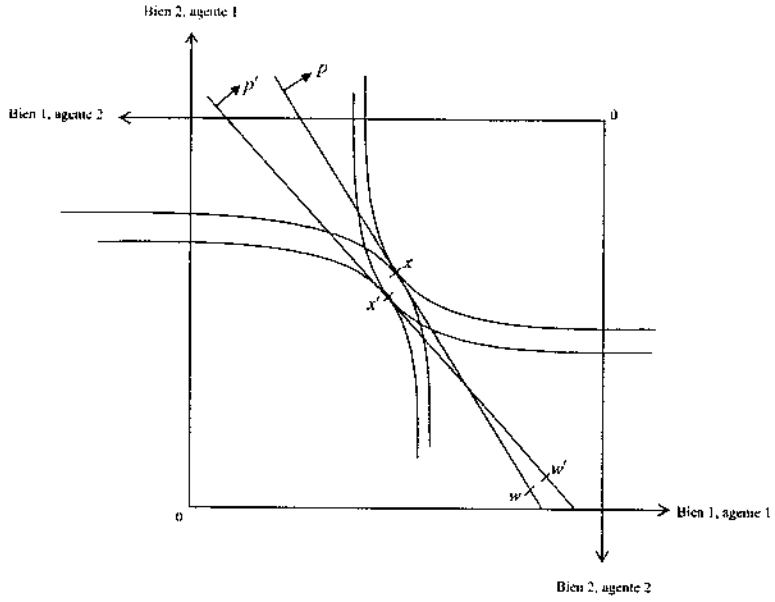
Donde se nota que en equilibrio el agente 1 es ofereute neto del bien 1 y demandante neto del bien 2.

Supongamos que, por alguna razón, la autoridad de política económica decide que debe mejorar el bienestar del agente 1. Por supuesto, la autoridad no puede aumentar la dotación agregada inicial de la economía (es decir, no puede agrandar la Caja de Edgeworth). La única herramienta de política disponible a la autoridad es la redistribución de las dotaciones iniciales. Supongamos que la autoridad realiza una política que, al menos inicialmente, parece adecuada: va a tomar de la dotación inicial del agente 2 cierta cantidad de cada bien y la va a transferir al agente 1. Al ser más rico, la autoridad espera que la situación del agente 1 sea mejor después de esta redistribución.

Wassily Leontief fue el primero en llamar la atención acerca de los riesgos de este razonamiento. El resultado fue posteriormente demostrado, en toda formalidad, por Marie-Paule Donsimoni y Herakles Polemarchakis. El argumento es el siguiente: supongamos que para el agente 1 el bien 1 es inferior y para el agente 2 este mismo bien es normal. Cuando se produce la transferencia de dotaciones, es decir a los precios p, la riqueza del agente 1 aumenta y la del agente 2 disminuye. Dado esto y nuestros supuestos acerca de las preferencias de los agentes, la demanda agregada por el bien 1 debe disminuir y, con ello, su precio relativo debe caer. ¡El problema es que el precio relativo del bien que el agente 1 ofrece a la economía está cayendo mientras que el de el bien que él demanda está aumentando! No hay ninguna razón para descartar la posibilidad de que a los nuevos precios la riqueza del agente 1 no sea suficiente para comprar la canasta que él demandaba en el equilibrio anterior a la implementación de la política.

En la gráfica, lo que estamos diciendo es que con las dotaciones w', podría ocurrir que el nuevo equilibrio

sea (p', x') en cuyo caso el agente 1 se encuentra peor que son el equilibrio de las dotaciones w.



ANEXO 4

FISICA CUANTICA
Átomos, Moléculas, Sólidos, Núcleos y Partículas. Editorial Limusa 1ra Edic.1993
Robert Eisberg - Robert Resnick

Radiación térmica y el postulado de Planck

1.1 Introducción

En una reunión de la Sociedad Alemana de Física, el 14 de diciembre de 1900, Max Planck leyó un trabajo intitulado "La teoría de la ley de distribución de energías del espectro normal". Este trabajo que en un principio atrajo poca atención, fue el precursor de una revolución en la física. La fecha de su presentación se considera como el nacimiento de la física cuántica, a pesar de que fue hasta un cuarto de siglo después, cuando Schrödinger y otros desarrollaron la mecánica cuántica moderna, base del conocimiento actual. Fueron muchos los caminos que convergieron en este conocimiento, cada uno de los cuales mostró distintos aspectos de las fallas de la física clásica. En éste y los siguientes capítulos se examinarán los logros más importantes de la que hoy se llama *teoría cuántica antigua* y que dio origen a la mecánica cuántica moderna. Los fenómenos experimentales que se analizarán en relación con la teoría cuántica antigua, comprenden todas las disciplinas de la física clásica: mecánica, termodinámica, mecánica estadística y electromagnetismo. La necesidad de una mecánica cuántica, se manifestará por la contradicción sistemática de las leyes clásicas respecto a dichos fenómenos y la solución a esos conflictos en base a ideas cuánticas. El estudio de la teoría cuántica antigua permitirá obtener, más fácilmente, un conocimiento más profundo de la mecánica cuántica cuando se inicie su consideración en el quinto capítulo.

Como en el caso de la relatividad (que se trata brevemente en el apéndice A), la física cuántica representa una generalización de la física clásica, que incluye a las leyes clásicas como casos particulares. Así como la relatividad extiende el campo de aplicación de las leyes de la física a la región de altas velocidades, la física cuántica lo extiende a la región de dimensiones pequeñas; y así como la relatividad se caracteriza por una constante universal de significado fundamental, la velocidad de la luz c, así mismo la física cuántica se caracteriza por una constante universal de significado fundamental, que hoy se llama constante de Planck h. En su trabajo de 1900, Planck introdujo esta constante para tratar de explicar las propiedades observadas en la radiación térmica. De esta manera se empezará el estudio de la radiación térmica, que conducirá a la constante de Planck y, relacionada con ésta, al concepto cuántico de energía discreta. También se verá que la radiación térmica en sí, es de gran importancia y actualidad, ya que, por ejemplo, este fenómeno ha ayudado recientemente a los astrofísicos a decidir entre varias teorías acerca del origen del universo.

1.2 Radiación térmica

Se llama *radiación térmica*, a la radiación emitida por un cuerpo como consecuencia de su temperatura. Todos los cuerpos emiten esta radiación a su derredor, y la absorben de él. Si, en

un principio, el cuerpo está más caliente que su alrededor, se enfriará, ya que la rapidez con que emite energía excederá la rapidez con que la absorbe. Cuando se alcanza el equilibrio térmico la rapidez de emisión y la de absorción de energía serán iguales. La materia en un estado condensado (es decir, sólido o líquido) emite un espectro de radiación continuo. Los detalles del espectro son casi independientes del material particular del cual se compone el cuerpo, pero dependen fuertemente de la temperatura. A temperaturas ordinarias, la mayoría de los cuerpos son visibles no por la luz que emiten sino por la luz que reflejan, ya que si no se hace incidir luz sobre ellos no es posible verlos. Sin embargo, a muy altas temperaturas, los cuerpos son luminosos por sí mismos. En un cuarto obscuro se les puede ver brillar; pero aún a temperaturas de varios miles de grados Kelvin, más del 90% de la radiación térmica emitida es invisible para nosotros, empezando por la parte correspondiente al infrarrojo del espectro electromagnético. Por lo tanto, los cuerpos luminosos por sí mismos están muy calientes.

Por ejemplo, considere el calentamiento en el fuego, de una barra de hierro a temperaturas cada vez más altas, retirando en forma periódica la barra del fuego, el tiempo suficiente para observar sus propiedades. Cuando la barra aún está a temperaturas relativamente bajas, radia calor pero no está visiblemente caliente; conforme aumenta la temperatura, la cantidad de radiación emitida por la barra aumenta muy rápidamente y empiezan a notarse efectos visibles. La barra empieza a verse de un color rojo opaco, después adquiere un color rojo brillante y, a muy altas temperaturas, un intenso color blanco azuloso. Es decir, a medida que aumenta la temperatura, el cuerpo emite más radiación térmica y la frecuencia de la radiación más intensa se vuelve cada vez mayor.

La relación que existe entre la temperatura de un cuerpo y el espectro de frecuencias de la radiación emitida, se utiliza en un dispositivo llamado pirómetro óptico. Este dispositivo es esencialmente un espectrómetro rudimentario que permite al operador estimar la temperatura de un cuerpo caliente, como una estrella, observando el color o la composición de frecuencias de la radiación térmica que emite. Existe un espectro continuo de radiación emitida, pero el ojo humano ve principalmente el color correspondiente a la emisión más intensa en la región visible. El sol, los filamentos de focos y carbones calientes, son ejemplos comunes de objetos que emiten radiación visible.

En términos generales, la forma detallada del espectro de radiación térmica emitida por un cuerpo caliente, depende de la composición del mismo. Sin embargo, experimentalmente se encuentra que solo hay una clase de cuerpos que emiten espectros térmicos de características universales. Estos son los llamados *cuerpos negros*, es decir, cuerpos en cuyas superficies absorben toda la radiación térmica que incide sobre ellos. El nombre resulta apropiado puesto que dichos cuerpos no reflejan luz y, por tanto, se ven negros. Un ejemplo de un (casi) cuerpo negro, sería *cualquier* objeto cubierto con una capa difusa de pigmento negro, como negro-bismuto o negro-humo. Más adelante se describió otro ejemplo bastante diferente. Independientemente de los detalles de su composición, *todos* los cuerpos negros a la misma temperatura emiten radiación térmica con el mismo espectro. Este hecho general se puede entender en base a argumentos clásicos que implican el equilibrio termodinámico. Sin embargo, la forma específica del espectro no puede obtenerse solamente de argumentos termodinámicos. Las propiedades universales de la radiación emitida por cuerpos negros los hacen objeto de no interés teórico especial y los físicos siempre trataron de encontrar una explicación a las características específicas de su espectro.

La distribución espectral de la radiación de un cuerpo negro se especifica por la cantidad $R_T(\nu)$, llamada *radiancia espectral*, definida tal que $R_T(\nu)\,d\nu$ es igual a la energía emitida en forma de radiación con frecuencias en el intervalo entre ν y $\nu + d\nu$ de un área unitaria de la superficie a temperatura absoluta T y por unidad de tiempo. En 1899, Lummer y Pringsheim realizaron una de las primeras mediciones precisas de esta cantidad. Utilizaron un instrumento esencialmente similar a los espectrómetros de prisma utilizados para medir espectros ópticos, excepto que para los prismas, lentes, etc, se utilizaron materiales especiales, de manera que

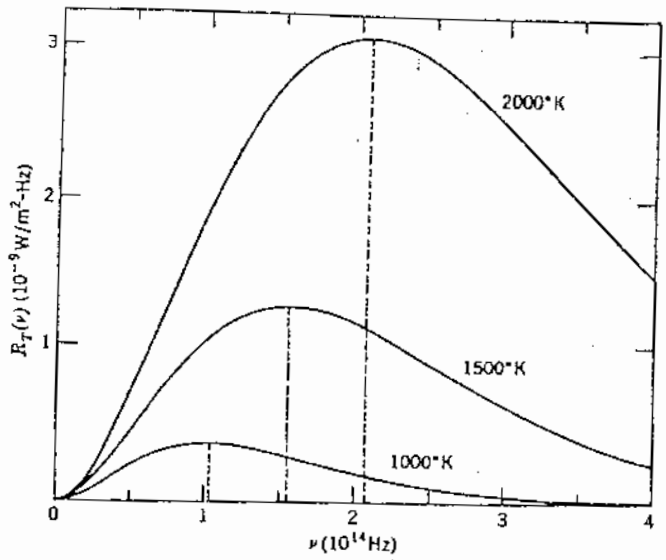


FIGURA 1-1

Radiancia espectral de un cuerpo negro radiante como función de la frecuencia de radiación, para temperaturas de 1000°K y 2000°K del cuerpo radiante. Obsérvese que la frecuencia a la que ocurre la máxima radiancia (línea punteada), aumenta linealmente conforme la temperatura aumenta y que la potencia total emitida, por metro cuadrado del cuerpo radiante (área bajo la curva), aumenta muy rápidamente con la temperatura.

fueran transparentes a la radiación térmica de frecuencia relativamente baja. En la figura 1-1 se indica la dependencia de $R_T(\nu)$ con T y ν que se observa en los experimentos.

Las *funciones de distribución*, de las cuales es un ejemplo, la radiancia espectral son muy comunes en física. Así por ejemplo, la función de distribución de velocidades de Maxwell (que se parece a una de las curvas de la figura 1-1) nos dice cómo se distribuyen las moléculas de un gas a presión y temperatura fijas, de acuerdo con su velocidad. Otra función de distribución, que probablemente el estudiante ha visto, es la que especifica el tiempo de decaimiento de núcleos radiactivos (que tiene la forma de una exponencial decreciente) en una muestra que contiene núcleos de una especie dada, y desde luego habrá visto la función de distribución correspondiente a las calificaciones obtenidas en un examen de física.

La función de distribución para la radiancia espectral de un cuerpo negro de área dada y a una temperatura particular, sea 1000°K de la figura 1-1, muestra que: (1) la potencia emitida por un intervalo de frecuencias pequeño $d\nu$, es pequeña si ese intervalo se encuentra a una frecuencia ν muy pequeña comparada con 10^{14} Hz. La potencia es cero para ν igual a cero. (2) La potencia radiada en el intervalo $d\nu$ aumenta rápidamente a medida que ν aumenta, partiendo de valores muy pequeños. (3) Alcanza un máximo para un valor de $\nu \simeq 1.1 \times 10^{14}$ Hz. Es decir, la potencia radiada con mayor intensidad, ocurre a esa frecuencia. (4) Por arriba de $\simeq 1.1 \times 10^{14}$ Hz, la potencia radiada disminuye lenta pero continuamente conforme ν aumenta. Se vuelve cero una vez más, cuando ν tiende a valores infinitamente grandes.

Las dos funciones de distribución, correspondientes a las temperaturas de 1500°K y 2000°K, que aparecen en la figura, muestran que: (5) la frecuencia para la cual la potencia radiada ocurre con mayor intensidad, aumenta, conforme la temperatura aumenta. Por inspección de la figura se puede verificar que esta frecuencia aumenta linealmente con la temperatura. (6) La potencia total radiada, en todas las frecuencias, aumenta conforme la temperatura aumenta, pero más rápidamente que en forma lineal. La potencia total radiada a una temperatura particular, se obtiene sencillamente del área bajo la curva correspondiente a esa temperatura. $\int_0^\infty R_T(\nu)\, d\nu$, ya que $R_T(\nu)d\nu$ es la potencia radiada en el intervalo de frecuencias comprendido entre ν y $\nu + d\nu$.

La integral de la radiancia espectral $R_T(\nu)$ sobre toda ν, es la energía total emitida de un cuerpo negro a temperatura T, por unidad de tiempo y por unidad de área. Se le llama la *radiancia* R_T, es decir,

$$R_T = \int_0 R_T(\nu)\, d\nu \tag{1-1}$$

Como se mencionó en el estudio relacionado con la figura 1-1, R_T aumentaría rápidamente a medida que la temperatura aumenta. El resultado se conoce como ley de Stefan, y fue enunciada por primera vez en 1879, en forma de una ecuación empírica:

$$R_T = \sigma T^4 \tag{1-2}$$

donde

$$\sigma = 5.67 \times 10^{-4}\ \mathrm{W/m^2\text{-}{}^{\circ}K^4}$$

es llamada *constante de Stefon-Boltzmann*. En la figura 1-1 también se muestra que el espectro se desplaza hacia frecuencias mayores a medida que T aumenta. Este resultado se conoce como *ley del desplazamiento de Wien*:

$$\nu_{\mathrm{máx}} \propto T \tag{1-3a}$$

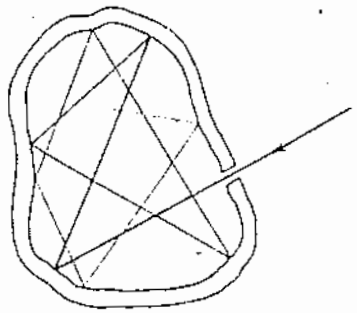
donde $\nu_{\mathrm{máx}}$ es lo frecuencia para la cual $R_T(\nu)$ alcanzo su valor máximo para una T particular. A medida que T aumenta, $\nu_{\mathrm{máx}}$ se desplaza hacia frecuencias mayores. Todos estos resultados concuerdan con las expériencias ya conocidas que se analizaron anteriormente, a saber; que la radiación térmica emitida aumenta rápidamente con la temperatura, (la barra de fierro radía mucha más energía térmica a temperaturas más altas), y la frecuencia principal de la radiación también aumenta, conforme la temperatura aumenta (la barra cambia de color, de rojo opaco a blanco azuloso).

Otro ejemplo de un cuerpo negro, que como se verá resulta particularmente importante, consiste en un objeto que contiene una cavidad y que se comunica con el exterior por medio de un pequeño agujero como se muestra en la figura 1-2. La radiación del exterior que incide sobre el agujero, penetra en la cavidad y se refleja hacia todos sentidos en las paredes de la cavidad, de modo que eventualmente se absorbe en estas paredes. Si el área del agujero es muy pequeña, comparada con el área de la superficie interna de la cavidad, la radiación reflejada hacia el exterior a través del agujero será despreciable. Esencialmente, toda la radiación que incide sobre el agujero será absorbida, por lo tanto, el *agujero* tendrá todas las propiedades de la superficie de un cuerpo negro. La mayoría de los cuerpos negros que se utilizan en los experimentos de laboratorio, se construyen a lo largo de estas líneas.

Supóngase que las paredes de la cavidad se calientan a una temperatura T, de modo que

FIGURA 1-2

Cavidad en un cuerpo comunicada con el exterior por medio de un pequeño agujero. La radiación incidente sobre el agujero es absorbida completamente después de reflexiones sucesivas en las paredes internas de la cavidad. El agujero absorbe radiación como un cuerpo negro. En el proceso inverso, por medio del cual la radiación que sale por el agujero se constituye por contribuciones emitidas de la superficie interna, el agujero emite radiación como un cuerpo negro.



emitirán radiación térmica que llenará la cavidad. Una pequeña fracción de esta radiación, que incida en el agujero, pasará por él, y así el agujero actuará como emisor de radiación térmica. Como el agujero debe tener las propiedades de la superficie de un cuerpo negro, la radiación que emite debe tener el espectro de un cuerpo negro; sin embargo, como el agujero simplemente muestrea la radiación térmica dentro de la cavidad, resulta evidente que la radiación en la cavidad también debe tener un espectro de cuerpo negro. De hecho, debe tener un espectro de cuerpo negro característico de la temperatura T en las paredes, ya que ésta es la única temperatura que se define en el sistema. El espectro emitido por el agujero en la cavidad, se especifica en términos del flujo de energía $R_T(\nu)$. Sin embargo, resulte más útil especificar el espectro de la radiación dentro de la cavidad, llamada *radiación de la cavidad*, en términos de una *densidad de energía*, $\rho_T(\nu)$, que se define como la energía contenida en una unidad de volumen de la cavidad a temperatura T, en el intervalo de frecuencia entre ν y $\nu + d\nu$. Es evidente que estas cantidades deben ser proporcionales entre sí, es decir,

$$\rho_T(\nu) \propto R_T(\nu) \tag{1-4}$$

Por lo tanto, la radiación dentro de una cavidad cuyas paredes están a temperatura T, tiene el mismo carácter que la radiación emitida por la superficie de un cuerpo negro a temperatura T. Experimentalmente, resulta conveniente producir un espectro de cuerpo negro, por medio de una cavidad en un cuerpo caliente con un agujero hacia el exterior, y teóricamente también es conveniente estudiar la radiación de un cuerpo negro, analizando la radiación de una cavidad, ya que es posible aplicar argumentos muy generales para predecir las propiedades de la radiación de una cavidad.

Ejemplo 1-1. (a) Ya que $\nu = c$, la velocidad constante de la luz, la ley del desplazamiento de Wien (1-3a), también se puede escribir como:

$$\lambda_{máx} T = \text{constante} \tag{1-3b}$$

donde $\lambda_{máx}$ es la longitud de onda para la cual, a una temperatura T particular, la radiancia espectral alcanza su valor máximo. El valor determinado experimentalmente para la constante de Wien es 2.898×10^{-3} m-°K. Si se supone que las superficies de las estrellas se comportan como cuerpos negros, se puede obtener una buena estimación de su temperatura, midiendo $\lambda_{máx}$. Para el sol, $\lambda_{máx} = 5100$ Å, mientras que para la estrella polar $\lambda_{máx} = 3500$ Å. Enumere la temperatura de la superficie de estas estrellas. (Un *angstrom* $= 1$ Å $= 10^{-10}$ m.)

Para el sol, $T = 2.898 \times 10^{-3}$ m·°K/5100×10^{-10} m $= 5700$°K. Para la estrella polar, $T = 2.898 \times 10^{-3}$ m-m·°K/3500×10^{-10} m $= 8300$°K.

A 5700°K, la superficie del sol está a una temperatura muy cercana a la necesaria para que la mayor parte de su radiación esté en la región visible del espectro. Lo anterior sugiere que durante las etapas de la evolución humana, nuestros ojos se han adaptado al sol, haciéndose más sensibles a aquellas longitudes de onda que emite con mayor intensidad.

b) Utilizando la ley de Stefan, (1-2), y las temperaturas recién obtenidas; determinar la potencia radiada por 1 cm² de superficie estelar.

Para el sol,

$$R_T = \sigma T^4 = 5.67 \times 10^{-8} \text{ W/m}^2\text{-°K}^4 \times (5700\text{°K})^4$$
$$= 5.90 \times 10^7 \text{ W/m}^2 \simeq 6000 \text{ W/cm}^2$$

Para la estrella polar,

$$R_T = \sigma T^4 = 5.67 \times 10^{-8} \text{ W/m}^2\text{-°K}^4 \times (8300\text{°K})^4$$
$$= 2.71 \times 10^8 \text{ W/m}^2 \simeq 27{,}000 \text{ W/cm}^2$$

Ejemplo 1-2. Suponga que se tienen dos cuerpos opacos, separados por una distancia grande y colgados de hilos finos en una cámara evacuada cuyas paredes se mantienen a temperatura constante. En tales condiciones, los cuerpos y las paredes intercambian calor sólo por medio de radiación. Sea e la razón de emisión de energía radiante de un cuerpo y sea a la razón de absorción de energía radiante del cuerpo. Demuestre que en equilibrio

$$\frac{e_1}{a_1} = \frac{e_2}{a_2} = 1 \tag{1-5}$$

Esta relación, (1-5), se conoce como *ley de radiación de Kirchhoff*. En todo el sistema dentro de la cámara y en tal estado, la razón de emisión necesariamente es igual a la razón de absorción para cada cuerpo. Entonces:

$$e_1 = a_1 \qquad \text{y} \qquad e_2 = a_2$$

Por lo tanto,

$$\frac{e_1}{a_1} = 1 = \frac{e_2}{a_2}$$

Si uno de los cuerpos, por ejemplo el cuerpo 2, es un cuerpo negro, entonces $a_2 > a_1$, ya que un cuerpo negro es más absorbente que uno que no es cuerpo negro. Por lo tanto, de (1-5) se sigue que $e_2 > e_1$. Así pues, el hecho observado de que cuerpos que son más absorbentes también son buenos emisores, se predice por la ley de Kirchhoff. ◀

1.3 Teoría clásica de la cavidad radiante

A principios del presente siglo, Rayleigh y también Jeans, hicieron cálculos de la densidad de energía de la radiación por una cavidad (o cuerpo negro), que señalaban hacia un serio conflicto entre la física clásica y los resultados experimentales. Este cálculo es similar a los que resultaron de considerar muchos otros fenómenos (es decir, el calor específico de los sólidos) que serán tratados más adelante. Los detalles serán presentados aquí, pero primero se describirán los procedimientos generales que servirán como guía en los cálculos.

Considérese que a una cavidad con paredes metálicas se la calienta uniformemente a una temperatura T. Las paredes emiten radiación electromagnética en el intervalo térmico de frecuencias. Se sabe que esto ocurre, básicamente, por el movimiento acelerado de los electrones en las paredes, que resulta de la agitación térmica (véase el apéndice B). Sin embargo, no es necesario estudiar en detalle el comportamiento de las ondas electromagnéticas en el interior de la cavidad. Rayleigh y Jeans procedieron en la forma siguiente. En primer lugar, se utiliza la teoría electromagnética clásica para demostrar que la radiación en el interior de la cavidad debe existir en forma de ondas estacionarias con nodos en las superficies metálicas. Utilizando argumentos geométricos, se cuenta el número de dichas ondas estacionarias en el intervalo de frecuencias entre ν y $\nu + d\nu$, con el fin de determinar *cómo* depende ese número de ν. Después, se utiliza un resultado de la teoría cinética para calcular la energía total promedio de estas ondas cuando el sistema está en equilibrio térmico. En la teoría clásica, la energía total promedio sólo depende de la temperatura T. El número de ondas estacionarias en el intervalo de frecuencias, multiplicado por la energía promedio de las ondas y dividido entre el volumen de la cavidad, proporciona el contenido de energía promedio por unidad de volumen en el intervalo de frecuencias entre ν y $\nu + d\nu$, y esta es la cantidad que se buscaba la densidad de energía $\rho_T(\nu)$. Ahora, haga todo esto.

Por simplicidad, supóngase que la cavidad de paredes metálicas llena con radiación electromagnética tiene la forma de un cubo de lado a, como se muestra en la figura 1-3. En ese caso, la radiación que se refleja de las paredes puede ser analizada en términos de tres componentes a lo

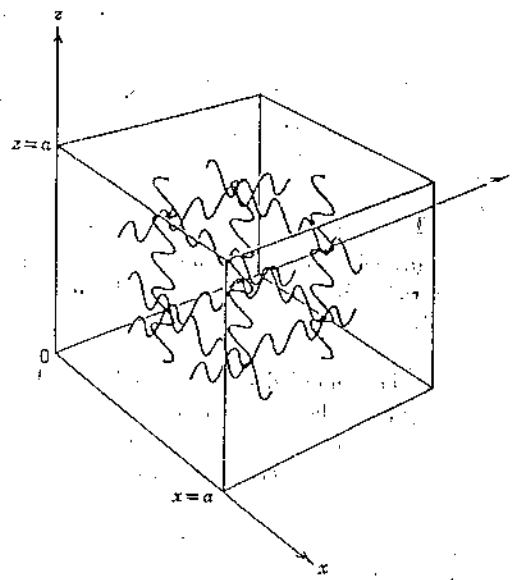


FIGURA 1-3

Cavidad con paredes metálicas llena con radiación electromagnética, mostrando las tres componentes de la radiación sin interferirse, que rebotan de las paredes y que forman ondas estacionarias con nodos en cada pared.

largo de las tres direcciones mutuamente perpendiculares que definen los lados de la cavidad. Como las paredes opuestas son paralelas entre sí, las tres componentes de la radiación no se mezclan y se pueden tratar por separado. Considérese la componente x en la pared metálica en $x = 0$. Toda la radiación que incide sobre esta pared es reflejada y las ondas incidentes y reflejadas se combinan para formar una onda estacionaria. Ahora bien, como la radiación electromagnética es una vibración trasversal con el vector de campo eléctrico E perpendicular a la dirección de propagación, y como la dirección de propagación para este componente es perpendicular a la pared en cuestión, el vector de campo eléctrico E es paralelo a la pared. Sin embargo, una pared metálica no es compatible con un campo eléctrico paralelo a su superficie, ya que siempre se puede establecer un flujo de cargas de modo tal que se neutralice el campo. Por lo tanto para esta componente, E siempre es cero en la pared. Es decir, la onda estacionaria asociada con la componente x de la radiación, debe tener un nodo (amplitud cero) en $x = 0$. La onda estacionaria también deberá tener un nodo en $x = a$, ya que no puede haber un campo eléctrico paralelo en la pared correspondiente. Además, se aplican condiciones similares a las otras dos componentes; la onda estacionaria asociada con la componente y deberá tener nodos en $y = 0$ y $y = a$, y la onda estacionaria asociada con la componente z, deberá tener nodos en $z = 0$ y $z = a$. Estas condiciones imponen limitaciones en las longitudes de onda posibles, y por lo tanto, en las frecuencias posibles de la radiación electromagnética en la cavidad.

Ahora se considerará el problema de contar el número de ondas estacionarias con nodos en las superficies de la cavidad, cuyas longitudes de onda se encuentran en el intervalo entre λ y $\lambda + d\lambda$, correspondiente al intervalo de frecuencias entre ν y $\nu + d\nu$. Primero se tratará la componente x, solamente para fijar la atención en las ideas contenidas en este cálculo; es decir, se tratará el caso simplificado, aunque artificial, de una "cavidad unidimensional" de longitud a. Al desarrollar este caso, se verá que la generalización al caso real tridimensional es obvio.

El campo eléctrico para una onda estacionaria unidimensional puede describirse matemáticamente por la función

$$E(x, t) = E_0 \operatorname{sen}(2\pi x/\lambda) \operatorname{sen}(2\pi\nu t) \qquad (1\text{-}6)$$

donde λ es la longitud de onda, ν la frecuencia y E_0 la amplitud máxima de la onda. Las dos primeras cantidades se relacionan entre sí por medio de la ecuación

$$\nu = c/\lambda \qquad (1\text{-}7)$$

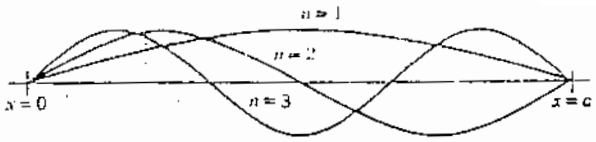


FIGURA 1-4

Patrones de amplitud para ondas estacionarias en una cavidad unidimensional con paredes en $x = 0$ y $x = a$, para los tres primeros valores del índice n.

donde c es la velocidad de propagación de las ondas electromagnéticas. La ecuación (1-6) representa una onda cuya amplitud varía en el espacio como sen $(2\pi x/\lambda)$ y que oscila en el tiempo con frecuencia ν como un oscilador armónico simple. Puesto que obviamente, la amplitud será igual a cero en todo tiempo para las posiciones que satisfagan la relación

$$2x/\lambda = 0, 1, 2, 3, \ldots \qquad (1\text{-}8)$$

la onda tendrá nodos fijos en ellas, esto es, es una onda estacionaria. Para poder satisfacer el requisito de que las ondas tengan nodos en los extremos de la cavidad unidimensional se escoge el origen del eje x en uno de los extremos de la cavidad ($x = 0$) y se exige que en el otro extremo ($x = a$)

$$2x/\lambda = n \qquad \text{para } x = a \qquad (1\text{-}9)$$

donde

$$n = 1, 2, 3, 4, \ldots$$

Esta condición determina el conjunto de valores permitidos de la longitud de onda. Para estos valores permitidos, las amplitudes de las ondas estacionarias siguen un patrón cuya apariencia se muestra en la figura 1-4. Estos patrones se pueden reconocer como los correspondientes a ondas estacionarias producidas por las vibraciones de una cuerda sujeta en sus extremos, sistema físico real que también satisface (1-6). En nuestro caso, el patrón representa ondas electromagnéticas estacionarias.

Por conveniencia, la discusión se continúa en términos de frecuencias permitidas en lugar de longitudes de onda permitidas. Estas frecuencias son $\nu = c/\lambda$, donde $2a/\lambda = n$. Es decir,

$$\nu = cn/2a \qquad n = 1, 2, 3, 4, \ldots \qquad (1\text{-}10)$$

Estos valores permitidos de la frecuencia pueden representarse por un diagrama que consiste de un eje sobre el cual se señalan puntos correspondientes a cada valor entero de n. En dicho diagrama, el valor permitido de la frecuencia ν, correspondiente a un valor particular de n, de acuerdo con (1-10), es igual a $c/2a$ veces la distancia d del origen al punto respectivo, o también, la distancia d es $2a/c$ veces la frecuencia ν. Estas relaciones se muestran en la figura 1-5. Dichos diagramas resultan útiles en el cálculo del número de valores permitidos de la frecuencia ν en el intervalo de ν a $\nu + d\nu$, que se denota por $N(\nu)\,d\nu$. Para evaluar esta cantidad, simplemente se cuenta el número de puntos sobre el eje n que se encuentran entre dos límites que se construyen de manera tal que correspondan a las frecuencias ν y $\nu + d\nu$, respectivamente, como los puntos se distribuyen uniformemente en el eje n, aparentemente el número de puntos comprendidos entre estos dos límites será proporcional a $d\nu$ pero no dependerá de ν. En realidad, fácilmente se ve que $N(\nu)\,d\nu = (2a/c)\,d\nu$. Sin embargo, lo anterior debe multiplicarse por un factor de 2, ya que, para cada frecuencia permitida, existen en realidad dos ondas independientes correspon-

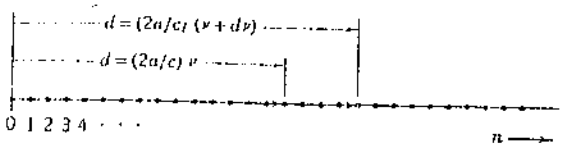


FIGURA 1-5

Valores permitidos del índice n, que determina los valores permitidos de la frecuencia, en una cavidad unidimensional de longitud a.

dientes a los dos estados de polarización posibles de las ondas electromagnéticas. Por lo tanto, se tiene que

$$N(\nu)\,d\nu = \frac{4a}{c}\,d\nu \tag{1-11}$$

De esta manera se completa el cálculo del número de ondas estacionarias permitidas para el caso artificial de una cavidad unidimensional.

El cálculo anterior hace evidente los procedimientos para extender el cálculo al caso real de una cavidad tridimensional. Esta extensión se indica en la figura 1-6. En este caso, el conjunto de puntos uniformemente distribuidos en valores enteros a lo largo del eje n, se sustituye por un arreglo tridimensional uniforme de puntos cuyas tres coordenadas corresponden a valores enteros a lo largo de tres ejes n mutuamente perpendiculares. *Cada punto* del arreglo corresponde a una onda estacionaria tridimensional particular permitida. El número de nodos de las componentes x, y, z de la onda estacionaria, se obtienen de los valores enteros n_x, n_y y n_z, equivale a analizar una onda tridimensional (es decir, que se propaga en una dirección arbitraria) en tres componentes tridimensionales. En este caso, el número de frecuencias permitidas en el intervalo de frecuencia de ν a $\nu = d\nu$ es igual al número de puntos contenidos entre dos cascarones esféricos con radios correspondientes a las frecuencias ν y $\nu + d\nu$ respectivamente.

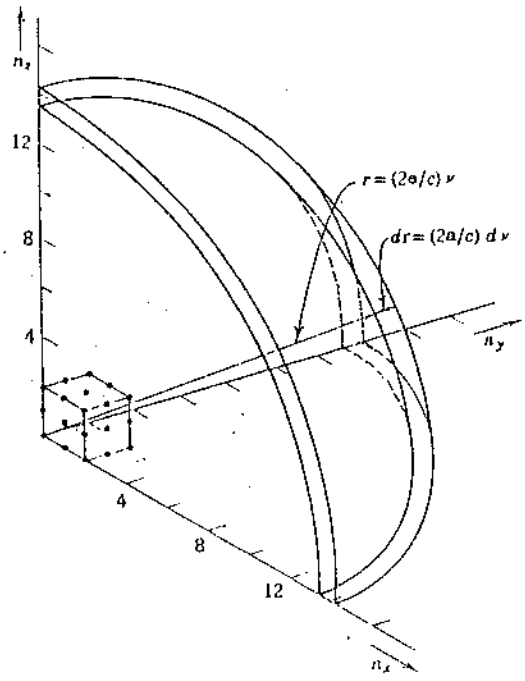


FIGURA 1-6

Los valores permitidos de la frecuencia en una cavidad tridimensional cúbica de lado a, se determinan por tres índices n_x, n_y, n_z que sólo pueden tomar valores enteros. Para mayor claridad, se muestran sólo unos cuantos de los muchos puntos correspondientes a conjuntos de estos tres índices.

Lo anterior será proporcional al volumen entre estos dos cascarones esféricos, ya que los puntos se distribuyen uniformemente. Por lo tanto, se puede ver que $N(\nu)\,d\nu$ será proporcional a $\nu^2\,d\nu$, ya que el factor ν^2 es proporcional al área de los cascarones y el segundo factor $d\nu$, es proporcional a la distancia entre ellos. En el ejemplo siguiente, se analizarán los detalles y se encontrará que

$$N(\nu)\,d\nu = \frac{8\pi V}{c^3}\,\nu^2\,d\nu \tag{1-12}$$

donde $V = a^3$, el volumen de la cavidad.

Ejemplo 1-3. Derivar (1-12), que dá el número de ondas electromagnéticas estacionarias permitidas, en cada intervalo de frecuencias, para el caso de una cavidad tridimensional cúbica de paredes metálicas de lado a.

Considérese radiación de longitud de onda λ y frecuencia $\nu = c/\lambda$, que se propaga en una dirección definida por los ángulos α, β, γ, como se muestra en la figura 1-7. La radiación debe ser una onda estacionaria ya que sus tres componentes son ondas estacionarias. La posición de algunos de los nodos fijos de esta onda estacionaria, se indica por un conjunto de planos perpendiculares a la dirección de propagación α, β, γ. La distancia entre estos planos nodales de la radiación, es justamente $\lambda/2$, donde λ es su longitud de onda. Se indica también la posición de los tres ejes de los nodos de las tres componentes.

La distancia entre estos nodos es:

$$\begin{aligned} \lambda_x/2 &= \lambda/2\cos\alpha \\ \lambda_y/2 &= \lambda/2\cos\beta \\ \lambda_z/2 &= \lambda/2\cos\gamma \end{aligned} \tag{1-13}$$

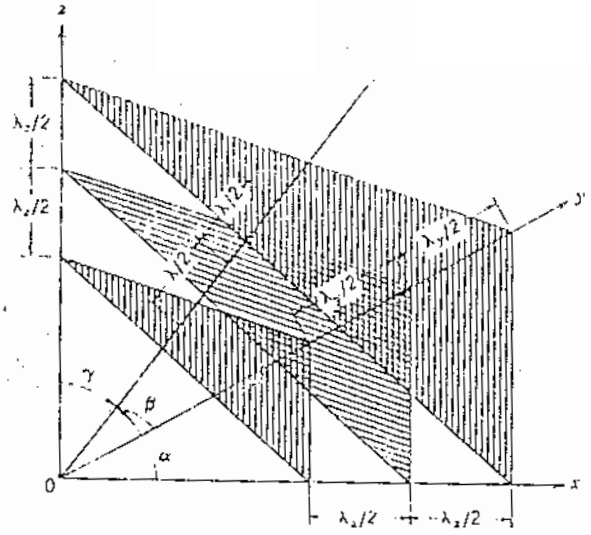


FIGURA 1-7

Planos nodales de una onda estacionaria propagándose en determinada dirección en una cavidad cúbica.

Las expresiones para las magnitudes de los campos eléctricos de las tres componentes en los tres ejes, se pueden escribir como,

$$E(x, t) = E_{0_x} \operatorname{sen}(2\pi x/\lambda_x) \operatorname{sen}(2\pi\nu t)$$
$$E(y, t) = E_{0_y} \operatorname{sen}(2\pi y/\lambda_y) \operatorname{sen}(2\pi\nu t)$$
$$E(z, t) = E_{0_z} \operatorname{sen}(2\pi z/\lambda_z) \operatorname{sen}(2\pi\nu t)$$

La expresión para la componente x representa una onda con amplitud máxima E_{0_x} con una variación en el espacio $\operatorname{sen}(2\pi x/\lambda_x)$ y que oscila con una frecuencia ν. Como el $\operatorname{sen}(2\pi x/\lambda_x)$ es cero para $2x/\lambda_x = 0, 1, 2, 3, \ldots$, la onda es estacionaria con longitud de onda λ_x ya que tiene nodos fijos separados una distancia $\Delta x = \lambda_x/2$. Las expresiones para las componentes y y z representan ondas estacionarias de amplitudes máximas E_{0_y}, E_{0_z} y longitudes de onda λ_y y λ_z pero las tres ondas estacionarias componentes oscilan con la frecuencia ν de la radiación. Obsérvese que las tres expresiones satisfacen automáticamente el requisito de que la componente x tenga un nodo en $x = 0$, la componente y tenga un nodo en $y = 0$, y la componente z tenga un nodo en $z = 0$. Para que satisfagan el requisito de que la componente x tenga un nodo en $x = a$, la componente y tenga un nodo en $y = a$, y la componente z tenga un nodo en $z = a$ se hace:

$$2x/\lambda_x = n_x \qquad \text{para } x = a$$
$$2y/\lambda_y = n_y \qquad \text{para } y = a$$
$$2z/\lambda_z = n_z \qquad \text{para } z = a$$

donde $n_x = 1, 2, 3, \ldots$; $n_y = 1, 2, 3, \ldots$; $n_z = 1, 2, 3, \ldots$. Usando (1-13), estas condiciones se transforman en:

$$(2a/\lambda)\cos\alpha = n_x \qquad (2a/\lambda)\cos\beta = n_y \qquad (2a/\lambda)\cos\gamma = n_z$$

Elevando al cuadrado ambos miembros de estas ecuaciones y sumando, se obtiene

$$(2a/\lambda)^2(\cos^2\alpha + \cos^2\beta + \cos^2\gamma) = n_x^2 + n_y^2 + n_z^2$$

pero los ángulos α, β, γ tienen la propiedad

$$\cos^2\alpha + \cos^2\beta + \cos^2\gamma = 1$$

Por lo tanto,

$$2a/\lambda = \sqrt{n_x^2 + n_y^2 + n_z^2}$$

donde n_x, n_y, n_z, toman todos los valores enteros posibles, esta ecuación describe la limitación en las longitudes de onda posibles de la radiación electromagnética contenida en la cavidad.

Una vez más, la discusión se continúa en términos de las frecuencias permitidas, en lugar de las longitudes de onda permitidas, que son

$$\nu = \frac{c}{\lambda} = \frac{c}{2a}\sqrt{n_x^2 + n_y^2 + n_z^2} \qquad (1\text{-}14a)$$

Enseguida, se contarán el número de frecuencias permitidas en un intervalo de frecuencias dado, construyendo una red cúbica uniforme en un octante de un sistema de coordenadas rectangulares, de modo tal, que las tres coordenadas de cada punto de la red son iguales a un conjunto posible de los tres enteros n_x, n_y y n_z, (ver la figura 1-6). Por construcción, cada punto de la red corresponde a una frecuencia permitida. Además, el número de frecuencias permitidas entre ν y $\nu + d\nu$, $N(\nu)\,d\nu$, es igual número de puntos contenidos entre cascarones esféricos concéntricos de radios r y $r + dr$, $N(r)\,dr$, donde

$$r = \sqrt{n_x^2 + n_y^2 + n_z^2}$$

De (1-14a), esto es

$$r = \frac{2a}{c}\nu \qquad (1\text{-}14b)$$

Como $N(r)\,dr$ es igual al volumen comprendido entre los cascarones por la densidad de puntos de la red, y como, por construcción, la densidad es uno, $N(r)\,dr$ es simplemente,

$$N(r)\,dr = \frac{1}{8}\,4\pi r^2\,dr = \frac{\pi r^2\,dr}{2} \tag{1-15}$$

Igualando lo anterior a $N(\nu)\,d\nu$, y evaluando $r^2\,dr$ de (1-14b), se tiene

$$N(\nu)\,d\nu = \frac{\pi}{2}\left(\frac{2a}{c}\right)^3 \nu^2\,d\nu$$

Para completar el cálculo, estos resultados se deben multiplicar por un factor de 2, ya que, para cada una de las frecuencias permitidas que se han enumerado, existen en realidad dos ondas independientes correspondientes a los dos estados de polarización posibles de la radiación electromagnética. Por lo tanto, se ha derivado (1-12). Se puede demostrar que $N(\nu)$ es independiente de la forma de la cavidad y sólo depende de su volumen. ◀

Obsérvese que existe una diferencia significativa entre los resultados obtenidos para el caso real de una cavidad tridimensional y los resultados obtenidos antes para el caso artificial de una cavidad unidimensional. El factor ν^2 que se encuentra en (1-12) pero no en (1-11), jugará un papel muy importante, como se verá, en los argumentos siguientes. Básicamente, este factor resulta del hecho de que vivimos en un mundo tridimensional siendo la potencia de ν uno menos que la dimensionalidad. Aunque Planck, al resolver finalmente las serias discrepancias entre la teoría clásica y el experimento, puso en tela de juicio algunos puntos que habían sido considerados obviamente ciertos, ni él ni otros trabajando sobre el problema dudaron de (1-12). Existió y aún persiste un acuerdo general de que (1-12) es válida.

Ahora ya se puede contar el número de ondas estacionarias. El paso siguiente en la teoría clásica de Rayleigh-Jeans para la radiación de un cuerpo negro, es el cálculo de la energía total promedio contenida en cada onda estacionaria de frecuencia ν. De acuerdo con la física clásica, la energía de alguna onda particular puede tener cualquier valor entre cero e infinito y su valor real debe ser proporcional al cuadrado de su amplitud constante E_0. Sin embargo, para un sistema que contenga un número grande de entes físicos del mismo tipo, los cuales están en equilibrio térmico a una temperatura T, la física clásica hace una predicción bastante definitiva acerca de los valores *promedio* de las energías de los entes. Esto puede aplicar en este caso, puesto que la multitud de ondas estacionarias, que constituyen la radiación térmica dentro de la cavidad, son entes del mismo tipo que están en equilibrio térmico entre sí, a la temperatura T de las paredes de la cavidad. El equilibrio térmico se asegura por el hecho de que las paredes de una cavidad real siempre absorberán y reirradiarán, con diferentes frecuencias y direcciones, una pequeña parte de la radiación que incide sobre ellas, y consecuentemente, las diferentes ondas estacionarias en forma gradual, podrán intercambiar energía como se requiere para mantener el equilibrio.

La predicción viene de la teoría cinética clásica, y es llamada *ley de equipartición de la energía*. Esta ley afirma que para un sistema de moléculas de un gas, en equilibrio térmico a una temperatura T, la energía cinética promedio de una molécula, por grado de libertad, es $kT/2$, donde $k = 1.38 \times 10^{-23}$ joule/°K se le llama *constante de Boltzmann*. Esta ley se aplica a cualquier sistema clásico que contenga, en equilibrio, un número grande de entes del mismo tipo. Para el caso que nos ocupa, los entes son ondas estacionarias que tienen como único grado de libertad, las amplitudes de sus campos eléctricos. Por lo tanto, sus energías cinéticas, en promedio, tendrán el mismo valor, $kT/2$. Sin embargo, cada onda estacionaria oscilante senoidal tiene una energía *total* que es el doble de su energía cinética promedio. Esta es una propiedad común a todos los sistemas que tienen un solo grado de libertad y que llevan a cabo oscilaciones armónicas simples en el tiempo; casos conocidos son un péndulo o un resorte. Por lo tanto, de

acuerdo con la ley clásica de la equipartición, cada onda estacionaria en la cavidad tiene una energía total promedio

$$\bar{\mathscr{E}} = kT \tag{1-16}$$

El punto más importante que debe notarse, es que se predice la misma energía total promedio $\bar{\mathscr{E}}$, para todas las ondas estacionarias en la cavidad, independientemente de sus frecuencias.

La energía por unidad de volumen en el intervalo de frecuencias entre ν y $\nu + d\nu$ del espectro de cuerpo negro de una cavidad a temperatura T, es simplemente el producto de la energía promedio de cada onda estacionaria, por el número de ondas estacionarias en el intervalo de frecuencias dividido entre el volumen de la cavidad. Finalmente, de (1-15) y (1-16) se obtiene el resultado,

$$\rho_T(\nu)\, d\nu = \frac{8\pi\nu^2 kT}{c^3}\, d\nu \tag{1-17}$$

Esta es la *fórmula de Rayleigh-Jeans, para la radiación de un cuerpo negro.*

En la figura 1-8 se comparan las predicciones de esta fórmula con los resultados experimentales. La discrepancia es evidente. En el límite de frecuencias bajas, el espectro clásico se aproxima a los resultados experimentales, pero, a medida que la frecuencia crece, la predicción teórica tiende a infinito. Los experimentos demuestran que la densidad de energía siempre permanece finita, como obviamente debe ser, y de hecho, la densidad de energía tiende a cero para frecuencias muy altas. Este comportamiento tan grotescamente irreal de la teoría clásica a altas frecuencias es conocido en física como la "catástrofe ultravioleta". Este término es sugestivo de la importancia de la falla de la teoría.

1.4 Teoría de Planck de la cavidad radiante

Tratando de resolver la discrepancia entre teoría y experimento, Planck llegó a considerar la posibilidad de que se violara la ley de equipartición de la energía en que se basaba la teoría. De la

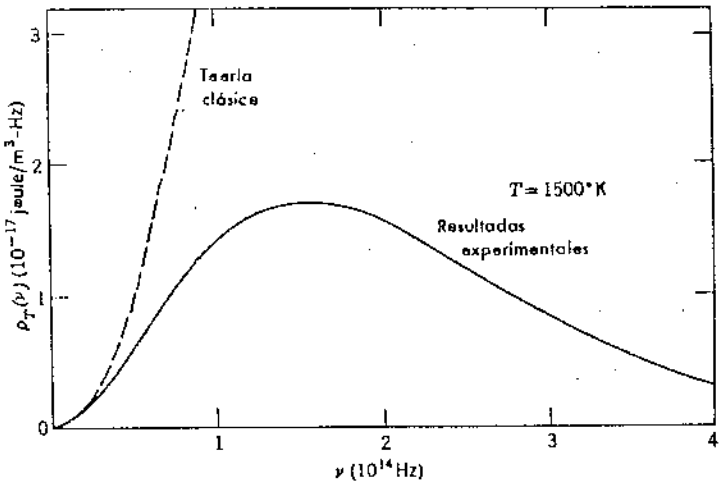


FIGURA 1-8

Predicción de Rayleigh-Jeans (línea punteada) comparada con los resultados experimentales (línea continua) para la densidad de energía en una cavidad-cuerpo negro que muestra la seria discrepancia llamada catástrofe ultravioleta.

figura 1-8 resulta evidente que la ley proporciona resultados satisfactorios a bajas frecuencia Por lo tanto se puede suponer

$$\bar{\mathscr{E}} \xrightarrow[\nu \to 0]{} kT \tag{1-1}$$

Es decir, la energía total promedio tiende a T a medida que la frecuencia tiende a cero. L discrepancia a frecuencias altas se elimina si, por alguna razón, existe un corte, de modo qu

$$\bar{\mathscr{E}} \xrightarrow[\nu \to \infty]{} 0 \tag{1-1}$$

es decir, si la energía total promedio tiende a cero cuando la frecuencia tiende a infinito. E otras palabras, Planck pensó que, dadas las circunstancias que prevalecen en el caso de radiación del cuerpo negro, la energía promedio de una onda estacionaria es función de frecuencia $\bar{\mathscr{E}}(\nu)$ con las propiedades indicadas por (1-18) y (1-19). Esto contrasta con la ley equipartición de la energía que asigna a la energía promedio $\bar{\mathscr{E}}$ un valor independiente de frecuencia.

Veamos el origen de la ley de equipartición. Básicamente surge de un resultado mucho m completo de la teoría cinética clásica llamado distribución de Boltzmann. (En el apéndice C dan los argumentos que conducen a la distribución de Boltzmann, para aquellos estudiantes q no estén familiarizados con ella.) Aquí se utilizará una forma especial de la distribución Boltzmann

$$P(\mathscr{E}) = \frac{e^{-\mathscr{E}/kT}}{kT} \tag{1-2}$$

en la que $P(\mathscr{E})\,d\mathscr{E}$ es la probabilidad de encontrar un ente dado de un sistema, con energía en intervalo entre $\mathscr{E}$ y $\mathscr{E} + d\mathscr{E}$, cuando el número de estados de energía para el ente en intervalo, es independiente de $\mathscr{E}$. Se supone que el sistema contiene un número grande de ent del mismo tipo, en equilibrio térmico a temperatura T, y k representa la constante Boltzmann. La ecuación (1-20) gobierna las energías de los entes del sistema que se est considerando, a saber, un conjunto de ondas estacionarias que oscilan de modo armóni simple.

La función de distribución de Boltzmann está íntimamente relacionada con la función distribución de Maxwell para la energía de una molécula en un sistema de moléculas equilibrio térmico. De hecho, la exponencial en la función de distribución de Boltzmann responsable por el factor exponencial en la distribución de Maxwell. Algunos estudian sabrán que existe un factor $\mathscr{E}^{1/2}$ presente en la distribución de Maxwell, que resulta por circunstancia de que el número de estados de energía para una molécula en el intervalo entre y $\mathscr{E} + d\mathscr{E}$ no es independiente de $\mathscr{E}$, sino que aumenta en proporción a $\mathscr{E}^{1/2}$.

La función de distribución de Boltzmann proporciona una información completa acerca las energías de los entes del sistema, incluyendo desde luego, el valor promedio de las energ $\bar{\mathscr{E}}$. Esta última cantidad puede obtenerse usando (1-20) para $P(\mathscr{E})$ y evaluando las integra en el cociente:

$$\bar{\mathscr{E}} = \frac{\int_0^\infty \mathscr{E} P(\mathscr{E})\,d\mathscr{E}}{\int_0^\infty P(\mathscr{E})\,d\mathscr{E}} \tag{1-}$$

El integrando en el numerador es la energía $\mathscr{E}$, pesada por la probabilidad de que el ente se encontrará con esta energía. La energía promedio se obtiene integrando sobre todos los valores posibles de la energía. El denominador es la probabilidad de encontrar al ente con *cualquier* energía y por lo tanto deberá tener el valor uno; que lo tiene. La integral en el numerador puede evaluarse, y el resultado es justamente la ley de equipartición de la energía

$$\bar{\mathscr{E}} = kT \tag{1-22}$$

En lugar de llevar a cabo la evaluación, resultará más conveniente, por los argumentos siguientes, examinar la representación gráfica de $P(\mathscr{E})$ y $\bar{\mathscr{E}}$ que se muestran en la parte superior de la figura 1-9. Se grafica $P(\mathscr{E})$ como función de $\mathscr{E}$. Su valor máximo, $1/kT$, ocurre en $\mathscr{E} = 0$ y el valor de $P(\mathscr{E})$ decrece suavemente a medida que $\mathscr{E}$ aumenta y tiende a cero cuando $\mathscr{E} \to \infty$. Es decir, el valor que es más probable que se encontrará en una medida de $\mathscr{E}$, será cero. Pero el promedio $\bar{\mathscr{E}}$ de los resultados que se encontrarían en cierto número de medidas de $\mathscr{E}$ es mayor que cero. La evaluación de $\bar{\mathscr{E}}$ a partir de $P(\mathscr{E})$, se indica en la mitad inferior de la figura 1-9.

La gran contribución de Planck surgió cuando pudo darse cuenta que podía lograr el corte requerido, indicado en (1-19), si modificaba el cálculo que conduce a $\bar{\mathscr{E}}$ a partir de $P(\mathscr{E})$, tratando la energía $\mathscr{E}$ como si fuera una *variable discreta* en lugar de la *variable continua* que definitivamente es desde el punto de vista de la física clásica. Esto puede hacerse, cuantitativamente, si se vuelve a escribir (1-21) en términos de sumas en lugar de integrales. Pronto se verá

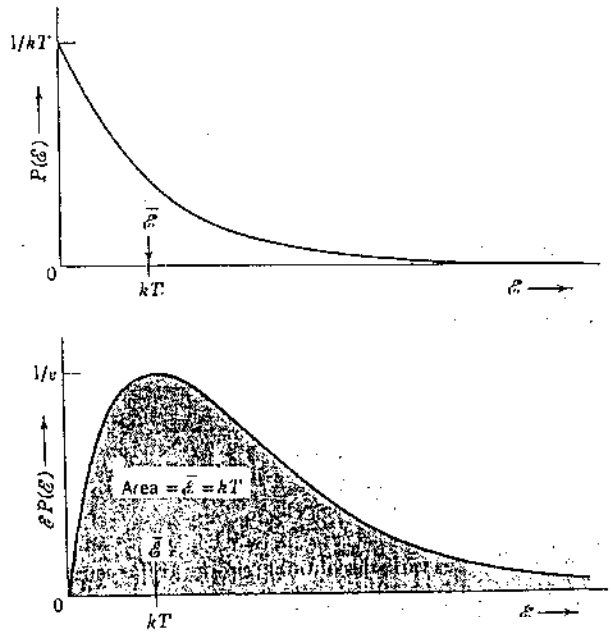


FIGURA 1-9

Arriba: Gráfica de la distribución de probabilidad de Baltzmann $P(\mathscr{E}) = e^{-\mathscr{E}/kT}/kT$. El valor promedio de la energía $\mathscr{E}$ para esta distribución es $\bar{\mathscr{E}} = kT$, que corresponde a la ley clásica de la equipartición de la energía. Para calcular este valor de $\bar{\mathscr{E}}$, se integra $\mathscr{E}P(\mathscr{E})$ de cero a infinito. Esta es justamente la cantidad que se está promediando, $\mathscr{E}$, multiplicada por la probabilidad relativa $P(\mathscr{E})$ de que el valor de $\mathscr{E}$ sea encontrado en una medida de la energía. *Abajo:* Gráfica de $\mathscr{E}P(\mathscr{E})$. El área bajo esta curva proporciona el valor de $\bar{\mathscr{E}}$.

que esto no es difícil de hacer, pero es más instructivo estudiar primero, la representación gráfica en la figura 1-10.

Planck supuso que la energía $\mathscr{E}$ podía tomar sólo ciertos valores discretos, en lugar de cualquier valor, y que los valores discretos de la energía estaban uniformemente distribuidos; es decir, tomó

$$\mathscr{E} = 0, \Delta\mathscr{E}, 2\Delta\mathscr{E}, 3\Delta\mathscr{E}, 4\Delta\mathscr{E}, \ldots \tag{1-23}$$

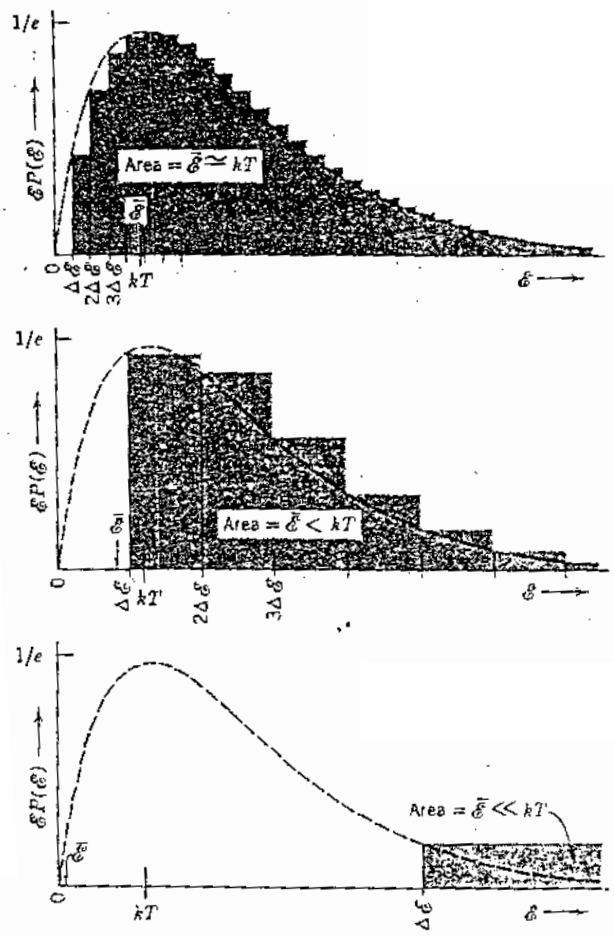


FIGURA 1-10

Arriba: Si la energía $\mathscr{E}$ no es una variable continua, sino que está restringida a valores $0, \Delta\mathscr{E}, 2\Delta\mathscr{E}, 3\Delta\mathscr{E}, \ldots$, como se indica por las divisiones en el eje $\mathscr{E}$ de la figura, se debe cambiar la integral usada para calcular el valor promedio $\bar{\mathscr{E}}$ por una suma. El valor promedio es la suma de las áreas de los rectángulos, de amplitud $\Delta\mathscr{E}$, y con alturas dadas por los valores permitidos de $\mathscr{E}$ por $P(\mathscr{E})$ al principiar cada intervalo. En la figura $\Delta\mathscr{E} \ll kT$, y por tanto, las energías permitidas están muy próximas entre sí y, consecuentemente, el área de los rectángulos difiere muy poco del área bajo la curva continua. Por lo tanto, el valor promedio $\bar{\mathscr{E}}$ es casi igual a kT, el valor encontrado en la figura 1-9. *Centro:* $\Delta\mathscr{E} \simeq kT$, y $\bar{\mathscr{E}}$ tiene un valor menor que el que tiene en el caso de arriba. *Abajo:* $\Delta\mathscr{E} \gg kT$, y $\bar{\mathscr{E}}$ es aún menos. En las tres figuras, los rectángulos muestran la contribución de $\mathscr{E}P(\mathscr{E})$ al área total, para cada energía permitida. El rectángulo para $\mathscr{E} = 0$ es, desde luego, siempre de altura cero. Esto produce un gran efecto en el área total, si las amplitudes de los rectángulos son grandes.

como el conjunto de valores permitidos de la energía. El intervalo uniforme entre valores sucesivos de la energía, es en este caso $\Delta\mathscr{E}$. En la parte superior de la figura 1-10, se ilustra una evaluación de $\bar{\mathscr{E}}$ a partir de $P(\mathscr{E})$, para un caso en el que $\Delta\mathscr{E} \ll kT$. En este caso el resultado que se obtiene es $\bar{\mathscr{E}} \simeq kT$. Es decir, el valor que aquí se obtiene es esencialmente el mismo que el resultado clásico, ya que, los valores discretos $\Delta\mathscr{E}$ son muy pequeños comparados con el intervalo de energías kT, en el cual $P(\mathscr{E})$ cambia significativamente; esencialmente, en este caso es indiferente si la energía $\mathscr{E}$ es discreta o continua. En la parte media de la figura 1-10, se ilustra el caso para $\Delta\mathscr{E} \simeq kT$. Aquí se encuentra que $\bar{\mathscr{E}} < kT$, ya que la mayoría de los entes tienen energía $\mathscr{E} = 0$, puesto que $P(\mathscr{E})$ tiene un valor pequeño para el primer valor permitido de $\Delta\mathscr{E}$ diferente de cero, de modo que el valor $\mathscr{E} = 0$ es dominante en el cálculo del valor promedio de $\mathscr{E}$ y el resultado que se obtiene es más pequeño. Sin embargo, en la parte inferior de la figura 1-10, donde se muestra el caso para $\Delta\mathscr{E} \gg kT$, se aprecia más claramente el efecto de tomar valores discretos. Para este caso, la probabilidad de encontrar un ente con una de las energías permitidas mayores que cero es despreciable, ya que, $P(\mathscr{E})$ es extremadamente pequeña para todos estos valores, y el resultado que se obtiene es $\bar{\mathscr{E}} \ll kT$.

Resumiendo, Planck descubrió que podía obtener $\bar{\mathscr{E}} \simeq kT$ cuando la diferencia entre energías adyacentes, $\Delta\mathscr{E}$, es pequeña y $\bar{\mathscr{E}} \simeq 0$, cuando $\Delta\mathscr{E}$ es grande. Como requería el primer resultado para valores pequeños de ν, y el segundo resultado para valores grandes de ν, obviamente era necesario que $\Delta\mathscr{E}$ fuese una función creciente de ν. Por trabajo numérico, encontró que podía tomar la relación más simple posible entre $\Delta\mathscr{E}$ y ν, que tuviera esta propiedad. Es decir, supuso que estas cantidades fueran proporcionales

$$\Delta\mathscr{E} \propto \nu \tag{1-24}$$

Escrita como una ecuación en lugar de una proporcionalidad, se tiene:

$$\Delta\mathscr{E} = h\nu \tag{1-25}$$

donde h es la constante de proporcionalidad.

Mediante trabajo numérico, Planck determinó el valor de esta constante, de modo que se obtuviera el mejor ajuste entre su teoría y los valores experimentales. El valor que obtuvo fue muy cercano al que es aceptado en la actualidad:

$$h = 6.63 \times 10^{-34} \text{ joule-seg}$$

Esta constante famosa, ahora se llama *constante de Planck*.

La fórmula que obtuvo Planck para $\bar{\mathscr{E}}$, evaluando la suma análoga a la integral en (1-21), y que se obtendrá en el ejemplo 1-4, es

$$\bar{\mathscr{E}}(\nu) = \frac{h\nu}{e^{h\nu/kT} - 1} \tag{1-26}$$

Ya que $e^{h\nu/kT} \to 1 + h\nu/kT$ para $h\nu/kT \to 0$, se tiene que $\bar{\mathscr{E}}(\nu) \to kT$ en este límite, como se predice en (1-18). En el límite $h\nu/kT \to \infty$, $e^{h\nu/kT} \to \infty$, y $\bar{\mathscr{E}}(\nu) \to 0$, en concordancia con la predicción de (1-19).

La fórmula que se obtiene inmediatamente para la densidad de energía en el espectro del cuerpo negro, utilizando el resultado para $\bar{\mathscr{E}}$ (ν) en lugar del valor clásico $\mathscr{E} = kT$, es:

$$\rho_T(\nu)\, d\nu = \frac{8\pi\nu^2}{c^3} \frac{h\nu}{e^{h\nu/kT} - 1}\, d\nu \tag{1-27}$$

Esto es, el *espectro del cuerpo negro de Planck*. En la figura 1-11, se comparan los resultados de la teoría de Planck (expresados en términos de longitudes de onda) con resultados experimentales

para una temperatura $T = 1595°K$. Los resultados experimentales concuerdan con la fórmula de Planck, para toda temperatura.

Debe recordarse que Planck no alteró la distribución de Boltzmann, "todo" lo que hizo fue considerar la energía de las ondas electromagnéticas estacionarias, oscilando senoidalmente en el tiempo, como una cantidad discreta en lugar de una cantidad continua.

Ejemplo 1-4. Derivar la expresión de Planck para la energía promedio $\bar{\mathscr{E}}$ y su espectro del cuerpo negro.

La cantidad $\bar{\mathscr{E}}$ se calcula de la razón de las sumas

$$\bar{\mathscr{E}} = \frac{\sum_{n=0}^{\infty} \mathscr{E} P(\mathscr{E})}{\sum_{n=0}^{\infty} P(\mathscr{E})}$$

análogas a la razón de integrales en (1-21). Debido al postulado de Planck, la energía $\mathscr{E}$ se convierte en una variable discreta que toma sólo los valores $\mathscr{E} = 0, h\nu, 2h\nu, 3h\nu, \ldots$ y por lo tanto, se deben utilizar sumas. Es decir, $\mathscr{E} = nh\nu$, donde $n = 0, 1, 2, 3, \ldots$. Evaluando la distribución de Boltzmann $P(\mathscr{E}) = e^{-\mathscr{E}/kT}/kT$, se tiene

$$\bar{\mathscr{E}} = \frac{\sum_{n=0}^{\infty} \frac{nh\nu}{kT} e^{-nh\nu/kT}}{\sum_{n=0}^{\infty} \frac{1}{kT} e^{-nh\nu/kT}} = kT \frac{\sum_{n=0}^{\infty} n\alpha e^{-n\alpha}}{\sum_{n=0}^{\infty} e^{-n\alpha}} \qquad \text{donde } \alpha = \frac{h\nu}{kT}$$

Lo anterior, a su vez, puede evaluarse más fácilmente si se observa que;

$$-\alpha \frac{d}{d\alpha} \ln \sum_{n=0}^{\infty} e^{-n\alpha} = \frac{-\alpha \frac{d}{d\alpha} \sum_{n=0}^{\infty} e^{-n\alpha}}{\sum_{n=0}^{\infty} e^{-n\alpha}} = \frac{-\sum_{n=0}^{\infty} \alpha \frac{d}{d\alpha} e^{-n\alpha}}{\sum_{n=0}^{\infty} e^{-n\alpha}} = \frac{\sum_{n=0}^{\infty} n\alpha e^{-n\alpha}}{\sum_{n=0}^{\infty} e^{-n\alpha}}$$

de manera que,

$$\bar{\mathscr{E}} = kT\left(-\alpha \frac{d}{d\alpha} \ln \sum_{n=0}^{\infty} e^{-n\alpha}\right) = -h\nu \frac{d}{d\alpha} \ln \sum_{n=0}^{\infty} e^{-n\alpha}$$

Ahora bien,

$$\sum_{n=0}^{\infty} e^{-n\alpha} = 1 + e^{-\alpha} + e^{-2\alpha} + e^{-3\alpha} + \cdots$$

$$= 1 + X + X^2 + X^3 + \cdots \qquad \text{donde } X = e^{-\alpha}$$

pero

$$(1 - X)^{-1} = 1 + X + X^2 + X^3 + \cdots$$

de modo que;

$$\bar{\mathscr{E}} = -h\nu \frac{d}{d\alpha} \ln (1 - e^{-\alpha})^{-1}$$

$$= \frac{-h\nu}{(1 - e^{-\alpha})^{-1}} (-1)(1 - e^{-\alpha})^{-2} e^{-\alpha}$$

$$= \frac{h\nu e^{-\alpha}}{1 - e^{-\alpha}} = \frac{h\nu}{e^{\alpha} - 1} = \frac{h\nu}{e^{h\nu/kT} - 1}$$

Se ha derivado (1-26) para la energía promedio de una onda electromagnética estacionaria de frecuencia ν. Multiplicando lo anterior por (1-12), a saber, el número $N(\nu)\,d\nu$ de ondas que tienen frecuencia ν, que se derivó en el ejemplo 1-3, inmediatamente se obtiene el espectro de Planck del cuerpo negro, (1-27). ◀

Ejemplo 1-5. Al analizar resultados experimentales, como en la figura 1-11, resulta conveniente expresar el espectro de Planck del cuerpo negro, en función de la longitud de onda λ, en lugar de la

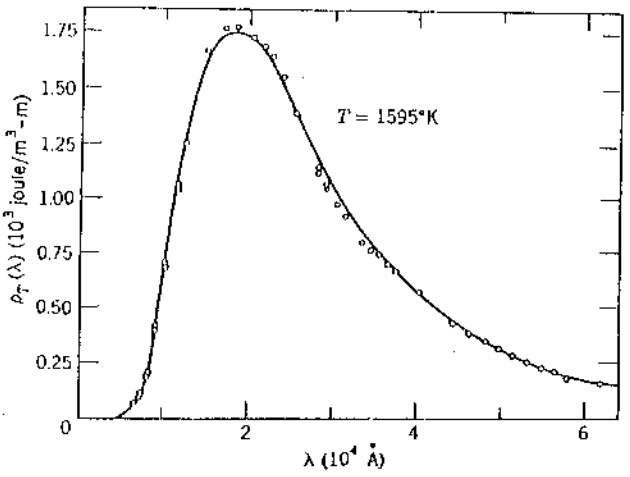


FIGURA 1-11

Predicción de Planck de la densidad de energía (línea continua) comparada con los resultados experimentales (círculos) para la densidad de energía de un cuerpo negra. Los datos fueron reportados por Cablentz en 1916, correspondientes a una temperatura de 1595°K. El autor señala en su artículo que después de dibujar las curvas de las espectros de energía que resultaron de sus experimentos, "debido a fatiga de los ojos, fue imposible durante los meses subsiguientes, dedicar atención a la reducción de los datos". Una vez que las datos fueran reducidos, se obtuvo un valor de la constante de Planck de 6.57×10^{-34} joule-seg.

frecuencia ν. Obtener $\rho_T(\lambda)$, el espectro de Planck en términos de la longitud de onda, a partir de $\rho_T(\nu)$, la forma del espectro en términos de ν. La cantidad $\rho_T(\lambda)$ se define por la igualdad $\rho_T(\lambda)\, d\lambda = -\rho_T(\nu)\, d\nu$. El signo menos indica que aunque $\rho_T(\lambda)$ y $\rho_T(\nu)$ son positivas ambas $d\lambda$ y $d\nu$ tienen signos opuestos. (Un aumento en la frecuencia corresponde a una disminución en la longitud de onda).

De la relación $\nu = c/\lambda$ se tiene que $d\nu = -(c/\lambda^2)\, d\lambda$, o bien, $d\nu/d\lambda = -(c/\lambda^2)$, de manera que

$$\rho_T(\lambda) = -\rho_T(\nu)\frac{d\nu}{d\lambda} = \rho_T(\nu)\frac{c}{\lambda^2}$$

Si ahora se sustituye $\nu = c/\lambda$ en la expresión para $\rho_T(\nu)$ (1-27), se obtiene

$$\rho_T(\lambda)\, d\lambda = \frac{8\pi hc}{\lambda^5}\frac{d\lambda}{e^{hc/\lambda kT} - 1} \tag{1-28}$$

En la figura 1-12 se muestra $\rho_T(\lambda)$ contra λ para diferentes temperaturas. Al estudiar la distribución de energía radiante como función de la longitud de onda, a medida que aumenta la temperatura, se puede observar, claramente, una tendencia en la radiación del roja al blanco y al azul. ◀

A partir de la fórmula de Planck, se pueden derivar la ley de Stefan y la ley del desplazamiento de Wien; si se las ajusta a los resultados experimentales, se pueden determinar valores para las constantes h y k. La ley de Stefan se obtiene integrando la ley de Planck en todo el espectro de longitudes de onda. Se encuentra que la radiancia es proporcional a la cuarta potencia de la temperatura, identificando la constante de proporcionalidad $2\pi^5 k^4/15c^2h^3$ con la constante de Stefan, σ, cuyo valor, determinada experimentalmente, es 5.67×10^{-8} W/m^2-°K^4. La ley del desplazamiento de Wien, se obtiene haciendo $d\rho(\lambda)/d\lambda = 0$, y se encuentra que $\lambda_{máx}T =$

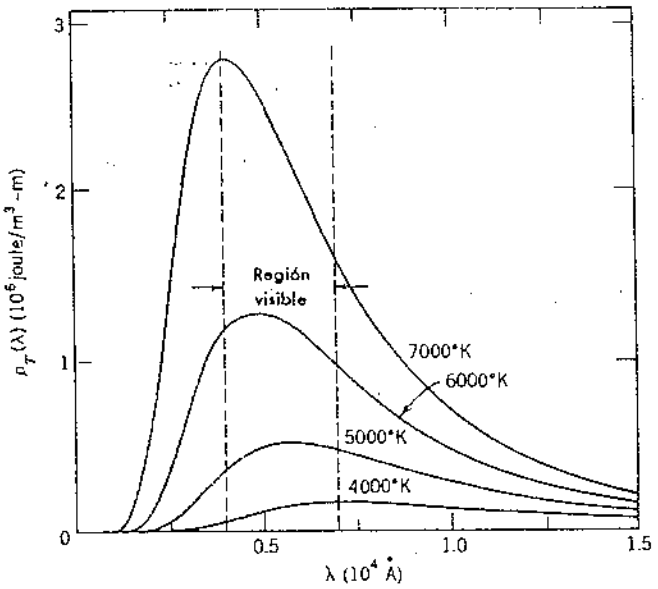


FIGURA 1-12

Densidad de energía de Planck de la radiación de un cuerpo negro, a varias temperaturas, como función de la longitud de onda. Obsérvese que la longitud de onda para la cual la curva tiene un máximo, decrese a medida que la temperatura aumenta.

$0.2014hc/k$; identificándose al miembro a la derecha de la ecuación con la constante, 2.898×10^{-3} m-°K, determinada experimentalmente por Wien. Utilizando estos valores medidos y suponiendo un valor para la velocidad de la luz c, se pueden calcular los valores de h y k. De hecho, ésto lo realizó Planck y sus valores concuerdan muy bien con los obtenidos posteriormente por otros métodos.

1.5 Aplicación de la ley de radiación de Planck en termometría

La radiación emitida por un cuerpo caliente puede utilizarse para medir su temperatura. Si lo que se utiliza es la radiación total, entonces, de la ley de Stefan-Boltzmann, se sabe que las energías emitidas por dos fuentes están en razón de la cuarta potencia de la temperatura. Sin embargo, es difícil medir la radiación total de la mayoría de las fuentes y realmente, lo que se mide es la radiancia en una banda finita de longitudes de onda. En este caso se utiliza la ley de radiación de Planck que proporciona la radiancia como función de la temperatura y longitud de onda. Para radiación monocromática de longitud de onda λ la razón de intensidades espectrales emitidas por fuentes a T_2°K y T_1°K, está dada por la ley de Planck como:

$$\frac{e^{hc/\lambda kT_1} - 1}{e^{hc/\lambda kT_2} - 1}$$

En esta expresión, si se toma T_1, como temperatura de referencia estándar, se puede determinar T_2 relativa a la estándar, midiendo esu razón en forma experimental. Este es el procedimiento empleado en la Escala Internacional de Temperaturas Prácticas, donde se utiliza el punto de fusión del oro, 1068°C, como punto fijo de referencia. Es decir, el *pirómetro óptico* estándar primario, está diseñado de modo tal, que compara la radiación de un cuerpo negro, a

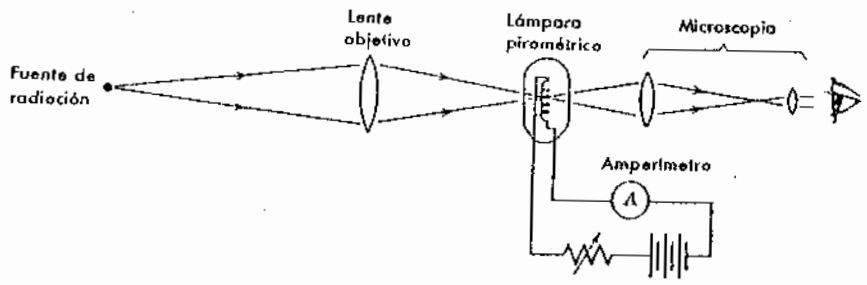


FIGURA 1-13

Diagrama esquemático de un pirómetro óptico.

temperatura desconocida $T > 1068°C$ con un cuerpo negro a la temperatura de fusión del oro. Dadas las circunstancias prácticas, la mayoría de las fuentes no son cuerpos negros y en lugar de la radiación monocromática se utiliza una banda finita de longitudes de onda, de modo que los procedimientos de medida deben ser adecuados y la teoría desarrollada, para tomar en consideración las circunstancias anteriores.

La mayoría de los pirómetros ópticos, utilizan el ojo humano como detector y requieren de un ancho de banda espectral suficientemente amplio, de modo que se tenga la energía que el ojo puede ver. El tipo de instrumento más simple y más preciso utilizado por encima del punto de fusión del oro, es el pirómetro óptico con filamento desvaneseente, (ver figura 1-13). La fuente a medirse se enfoca en el filamento de la lámpara del pirómetro, ajustándose la corriente de la lámpara hasta que el filamento tienda a desaparecer en el fondo de la imagen de la fuente. Calibrando cuidadosamente mediante potenciómetros de precisión, se garantizan medidas precisas de la temperatura.

En la década 1950, Dicke, Penzias y Wilson descubrieron un interesante caso particular del campo general de la termometría por medio de radiación de cuerpo negro. Utilizando un radio telescopio que operaba en el intervalo de longitudes de onda entre varios milímetros y varios centímetros, encontraron que un espectro de cuerpo negro, con temperatura característica de 3°K, incide sobre la tierra en todas direcciones con igual intensidad. La uniformidad en dirección indica que la radiación llena uniformemente el universo. Los astrofísicos consideraron estas medidas como evidencia fuerte en favor de la llamada *teoría de la gran explosión*, en la cual, el universo consistía de una bola de fuego muy densa y muy caliente de radiación y partículas, aproximadamente hace 10^{10} años. Debido a una expansión posterior y el consecuente corrimiento Doppler, es de esperarse que la temperatura de la radiación disminuyera a un valor muy cercano del observado de 3°K.

1.6 El postulado de Planck y sus implicaciones

La contribución de Planck se puede enunciar en forma de postulado, como sigue:

Cualquier ente físico con un grado de libertad, cuya "coordenada" es una función senoidal del tiempo (es decir realiza oscilaciones armónico-simples) sólo puede poseer energías totales $\mathscr{E}$, que satisfacen la relación

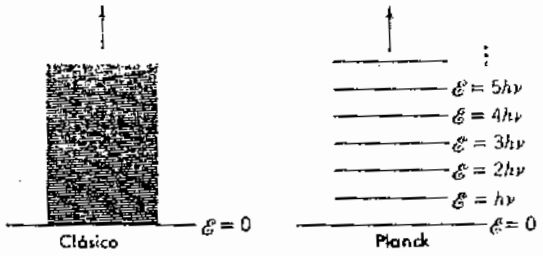
$$\mathscr{E} = nh\nu \qquad n = 0, 1, 2, 3, \ldots$$

donde ν es la frecuencia de la oscilación y h es una constante universal.

La palabra coordenada se utiliza en su sentido más general, para denotar cualquier cantidad que describa la condición instantánea del ente, por ejemplo, la longitud de un resorte, la

FIGURA 1-14

Izquierda: Energías permitidas en un sistema clásico, oscilando senoidalmente con frecuencia ν, distribuidas en forma continua. *Derecha:* Energías permitidas distribuidas en forma discreta de acuerdo con el postulado de Planck, ya que sólo pueden tener los valores $nh\nu$. Se dice que la energía está cuantizada, siendo n el número cuántico del estado de energía permitido.



posición angular de la masa de un péndulo y la amplitud de una onda. Todos estos ejemplos también son funciones senoidales del tiempo.

Un modo conveniente de ilustrar el comportamiento de un ente gobernado por este postulado, es mediante un diagrama de niveles de energía, como se muestra en la figura 1-14, que también puede utilizarse para contrastar este comportamiento con el que se esperaría en base a la física clásica. En ese diagrama se indican cada uno de los estados de energía posibles del ente, por una línea horizontal. La distancia de una línea determinada, a la línea de energía cero, es proporcional a la energía total que le corresponde. Como de acuerdo con la física clásica, el ente puede tener cualquier energía entre cero e infinito, el diagrama clásico de niveles de energías consiste de un continuo de líneas que empiezan en cero y se extiende a valores mayores. Sin embargo, el ente que realiza oscilaciones del tipo oscilador armónico simple, sólo puede tener energías totales discretas, $\mathscr{E} = 0, h\nu, 2h\nu, 3h\nu \ldots$, si obedece el postulado de Planck. En su diagrama de niveles de energía, esto se indica con un conjunto discreto de líneas. Si la energía del ente obedece el postulado de Planck, se dice que está *cuantizada*, los niveles de energía permitidos se llaman *estados cuánticos* y el número entero n se llama el *número cuántico*.

El estudiante podría pensar que existen sistemas físicos cuyo comportamiento esté en obvio desacuerdo con el postulado de Planck. Por ejemplo, un péndulo ordinario realiza oscilaciones de tipo oscilador armónico simple y sin embargo, aparentemente es capaz de poseer valores continuos de la energía. Pero antes de aceptar este argumento, se harán algunos cálculos numéricos en relación con dicho sistema.

Ejemplo 1-6. Un péndulo que consiste de una masa de 0.01 Kg. suspendida de una cuerda de 0.1 m. de longitud. Sea la amplitud de su oscilación tal, que la cuerda en sus posiciones extremas forma un ángulo de 0.1 rad. con la vertical. La energía del péndulo decrece, por ejemplo, por efectos de fricción. ¿Se observará que este decremento de la energía es continuo o discontinuo?

La frecuencia de oscilación del péndulo es:

$$\nu = \frac{1}{2\pi}\sqrt{\frac{g}{l}} = \frac{1}{2\pi}\sqrt{\frac{9.8 \text{ m/seg}^2}{0.1 \text{ m}}} = 1.6/\text{seg}$$

La energía del péndulo es su energía potencial máxima

$$mgh = mgl(1 - \cos\theta) = 0.01 \text{ kg} \times 9.8 \text{ m/seg}^2 \times 0.1 \text{ m} \times (1 - \cos 0.1)$$
$$= 5 \times 10^{-5} \text{ joule}$$

La energía del péndulo está cuantizada, de modo que los cambios en energía ocurren en pasos de magnitud discontinuos $\Delta E = h\nu$, pero

$$\Delta E = h\nu = 6.63 \times 10^{-34} \text{ joule-seg} \times 1.6/\text{seg} = 10^{-33} \text{ joule}$$

mientras que $E = 5 \times 10^{-5}$ joule. Por lo tanto, $\Delta E/E = 2 \times 10^{-29}$. Así pues, para poder medir que el decremento de la energía es discreto, se necesita poder medir la energía con una precisión de 2 en 10^{29}.

Evidentemente aún el equipo experimental más sensible, resulta totalmente incapaz para obtener dicha resolución en la energía. ◀

En conclusión, aquellos experimentos que involucren péndulos ordinarios, no podrán determinar si el postulado de Planck es válido o no. La pequeñez de h, hace que la granularidad de la energía sea tan fina que no pueda distinguirse de un continuo de energías. Más aún, para sistemas clásicos, la h podría considerarse igual a cero y, de hecho, un modo de reducir las fórmulas cuánticas a sus límites clásicos, es dejando $h \to 0$ en estas fórmulas. Solamente cuando se consideren sistemas en los cuales ν sea tan grande y/o $\mathscr{E}$ tan pequeña que $\Delta\mathscr{E} = h\nu$ sea del orden de $\mathscr{E}$, se estará en condiciones de probar el postulado de Planck. Un ejemplo es, desde luego, la radiación de ondas estacionarias de alta frecuencia del cuerpo negro. En los siguientes capítulos se considerarán muchos otros ejemplos.

1.7 Un poco de la historia del quantum

El postulado de Planck, en su forma original, no tenía el alcance que tiene en la forma como aquí se ha presentada. En su trabajo original, Planck trató en detalle, el comportamiento de los electrones en las paredes de un cuerpo negro y su acoplamiento a la radiación electromagnética dentro de la cavidad. Este acoplamiento conduce al mismo factor ν^2 que se obtuvo en (1-12) de argumentos más generales debidos a Rayleigh y Jeans. A través de este acoplamiento, Planck relacionaba la energía en una componente particular de las frecuencias de radiación del cuerpo negro, con la energía de un electrón en la pared, oscilando sencillamente a la misma frecuencia y postuló que únicamente la energía de la partícula oscilante está cuantizada. No fue sino hasta un tiempo después, que, Planck aceptó la idea de que las ondas electromagnéticas oscilantes estuvieran así mismo cuantizadas y el postulado se amplió para incluir cualquier ente cuya única coordenada oscile senoidalmente.

En un principio, Planck no estaba seguro si su introducción de la constante h no era más que un recurso matemático a un asunto con significado física muy profunda. En una carta dirigida a R.W. Wood, Planck llamó a su postulado limitado, "un acto de desesperación". "Sabía", escribió, "que el problema (del equilibrio de la materia con la radiación) tiene una importancia fundamental para la física; conocía la fórmula que reproduce la distribución de energías en el espectro normal; una interpretación teórica tenía que encontrarse a cualquier precio, sin importar qué tan alto". Por más de una década, Planck trató de ajustar las ideas cuánticas a la teoría clásica. Con cada intento aparentaba retroceder de su osadía original, pero siempre generaba nuevas ideas y técnicas que posteriormente adaptó la teoría cuántica. Lo que parece ser que finalmente lo convenció de que su hipótesis cuántica era correcta y de significado profundo, fue su apoyo a la definitividad del concepto estadístico de la entropía y a la tercera ley de la termodinámica.

Fue durante este período de dudas, que Planck era editor de la revista de investigación alemana *Annalen der Physik*. En 1905 recibió el primer artículo de Einstein sobre relatividad, y valientemente defendió el trabajo de Einstein. A partir de entonces se convirtió en una de los defensores del joven Einstein en los círculos científicos, pero resistió durante algún tiempo, precisamente aquellas ideas de Einstein, sobre la teoría cuántica de la radiación, que contribuyeron a confirmar y extender el propio trabajo de Planck. La profunda visión que Einstein tenía del electromagnetismo y la mecánica estadística, que posiblemente no fue igualada por sus contemporáneos, le permitió predecir, como resultado del trabajo de Planck, la necesidad de un cambio general en la mecánica estadística y el electromagnetismo. Muchos fenómenos físicos fueron predichos e interpretados anticipadamente las cuales posteriormente fueron dramáticamente confirmados experimentalmente. En el siguiente capítulo se tratará uno de estos fenómenos y se seguirá otro camino rumbo a la mecánica cuántica.

PREGUNTAS

1. ¿Un cuerpo negro siempre se ve negro? Explique el término cuerpo negro.

2. En un fuego hecho con carbón, los huecos que quedan entre carbones se ven más brillantes que los

carbones mismos. ¿Será la temperatura de dichos huecos, apreciablemente mayor que la correspondiente a la superficie exterior de un carbón incandescente?

3. Si se observa dentro de una cavidad cuyas paredes se mantienen a una temperatura constante, ¿es visible algún detalle del interior? Explíquese.

4. La relación $R_T = \sigma T^4$ es exacta para cuerpos negros y se mantiene a toda temperatura. ¿Por qué no se utiliza esta relación como base para una definición de temperatura, por ejemplo, 100°C?

5. Un pedazo de metal se pone incandescente con un color rojo brillante a 1100°K. Sin embargo, a esta misma temperatura, un pedazo de cuarzo simplemente no brilla. Explique. (Sugerencia: el cuarzo es transparente a la luz visible).

6. Haga una lista de las funciones de distribución que más comúnmente se usan en las ciencias sociales (es decir, distribución de familias respecto de sus ingresos). En cada caso, especifique si la variable de la distribución descrita es discreta o continua.

7. En (1-4) se relacionan radiancia espectral y densidad de energía, ¿qué unidades deberá tener la constante de proporcionalidad?

8. ¿Cuál es el origen de la catástrofe ultravioleta?

9. La ley de equipartición de la energía requiere que el calor específico de los gases sea independiente de la temperatura, en desacuerdo con la experimentación. Aquí se ha visto que conduce a la ley de radiación de Rayleigh-Jeans, que también está en desacuerdo con la experimentación. ¿se pueden relacionar estas dos fallas de la ley de equipartición?

10. Comparar las definiciones y dimensiones de radiancia espectral, $R_T(\nu)$, radiancia, R_T, y densidad de energía $\rho_T(\nu)$.

11. ¿Por qué se utiliza la pirometría óptica por arriba del punto de fusión del oro y no por abajo de él? ¿En qué objetos se mide típicamente la temperatura de este modo?

12. ¿Existen cantidades cuantizadas en física clásica? ¿Está cuantizada la energía en física clásica?

13. ¿Tiene sentido hablar en física de cuantización de carga? ¿En qué sentido difiere de la cuantización de la energía?

14. Las partículas elementales parecen tener un conjunto discreto de masas en reposo. ¿Se puede ésto considerar como cuantización de masa?

15. En muchos sistemas clásicos las frecuencias permitidas están cuantizadas. Mencione algunos de dichos sistemas. ¿También se cuantiza la energía?

16. Demuestre que la constante de Planck tiene dimensiones de impulso angular. ¿Sugiere ésto, necesariamente, que el impulso angular es una cantidad cuantizada?

17. ¿Cuál deberá ser el orden de magnitud mínimo de h, para que los efectos cuánticos sean efectos cotidianos?

18. La radiación universal de cuerpo negro a 3°K, ¿qué puede decir, si acaso, de la temperatura del espacio exterior?

19. La teoría de Planck ¿sugiere estados de energía atómicos cuantizados?

20. Discutir el hecho sorprendente de que la discresión de la energía se descubrió, por primera vez, al analizar el espectro continuo emitido por átomos interactuantes en un sólido, en lugar de encontrarlo en el análisis de un espectro discreto tal como el emitido por un átomo aislado en un gas.

PROBLEMAS

1. ¿A qué longitud de onda, una cavidad a 6000°K radiará más por unidad de longitud de onda?

2. Demuestre que la constante de proporcionalidad en (1-4) es $4/C$. Es decir, demuestre que la relación entre la radiancia espectral $R_T(\nu)$ y la densidad de energía $\rho_T(\nu)$ es $R_T(\nu)\,d\nu = (c/4)\rho_T(\nu)\,d\nu$.

3. Considere dos cavidades de forma y material arbitrarias, cada una a la misma temperatura T, conectadas por un tubo angosto en el que se pueden colocar filtros de colores (supuestamente ideales) que permitirán el paso de radiación de una sola frecuencia ν. (a) Suponga que para cierta frecuencia ν', $\rho_T(\nu')\,d\nu$ para la cavidad 1, resulta mayor que $\rho_T(\nu')\,d\nu$ para la cavidad 2. Se coloca un filtro en el tubo interconector, que sólo deja pasar la frecuencia ν. Analice qué sucederá en términos de flujo de energía. (b) Qué pasará con sus temperaturas respectivas. (c) Demuestre que esta violaría la segunda ley de la termodinámica; por lo tanto, pruebe que todos los cuerpos negros a la misma temperatura emitirán radiación térmica con el mismo espectro, independientemente de su composición.

4. Un radiador de cavidad, tiene un agujero de 10.0 mm. de diámetro que se perforó en una pared. Encuentre la potencia radiada a través del agujero, en el intervalo entre 5500 - 5510 Å. (Sugerencia: ver el problema 2).

5. (a) Suponiendo que la temperatura en la superficie del sol es 5700°K, utilice la ley de Stefan, (1-2), para determinar la masa en reposo que se pierde por segundo en la radiación del sol. Tome el diámetro del sol como 1.4×10^9 m. (b) ¿Qué fracción de la masa en reposo del sol, se pierde cada año en radiación electromagnética? Suponga que la masa en reposo del sol es 2.0×10^{30}kg.

6. En una explosión termonuclear, la temperatura en la bola de fuego es, momentáneamente, 10^7 °K. Encuentre la longitud de onda para la cual la radiación emitida es máxima.

7. A una temperatura dada $\lambda_{máx} = 6500$ Å para una cavidad de cuerpo negro. ¿Cuál será $\lambda_{máx}$ si la temperatura de las paredes de la cavidad aumenta de modo que la razón de emisión de radiación espectral se duplica?

8. ¿A qué longitud de onda emite el cuerpo humano su radiación térmica máxima? Haga una lista de las hipótesis que se hagan para llegar a una respuesta.

9. Suponiendo que $\lambda_{máx}$ está en el infrarrojo cercano para calor rojo, y en el ultravioleta cercano para calor azul. En la ley del desplazamiento de Wien aproximadamente, ¿qué temperatura corresponderá al calor rojo? ¿Al calor azul?

10. La razón *promedio* de radiación solar que incide sobre la tierra, por unidad de área es 0.485 cal/cm²-min (o 355 W/m²). (a) Explique la consistencia de este número con la constante solar (energía solar que llega a la tierra por unidad de tiempo incidiendo normal a una unidad de área de la superficie terrestre) cuyo valor es 1.94 cal/cm²-min (o 1340 W/m²). (b) Considere la tierra como un cuerpo negro que radia energía al espacio en esta misma razón. ¿Cuál sería la temperatura de la superficie de la tierra bajo estas circunstancias?

11. Demuestre que la ley de radiación de Rayleigh-Jeans, (1-17), no es consistente con la ley del desplazamiento de Wien $\nu_{max} \propto T$, (1-3a), o $\lambda_{max} T = \text{const}$, (1-3b).

12. A partir del espectro del cuerpo negra, ν_{max} se puede obtener de $d\rho_T(\nu)/d\nu = 0$ y λ_{max} de $d\rho_T(\lambda)/d\lambda = 0$. ¿Por qué a partir de $\lambda_{máx}\,T = \text{const}$ o $\nu_{max} = \text{const} \times T$ no es posible obtenerlas si se utiliza simplemente $\lambda_{max} = c/\nu_{max}$? Es decir, ¿por qué es erróneo suponer que $\nu_{max}\lambda_{max} = c$, *donde* c es la velocidad de la luz?

13. Considere los números siguientes: 2, 3, 3, 4, 1, 2, 2, 1, 0, como representando el número de cuadrangulares, logrado por cada miembro de los Orioles de Baltimore en un juego reciente. (a) Calcular directamente el número promedio de cuadrangulares por hombre. (b) Sea x una variable que

especifica el número de cuadrangulares obtenidos por un hombre, y sea $f(x)$ el número de veces que el número x aparece. Demuestre que el número promedio de cuadrangulares por hombre se puede escribir como

$$\bar{x} = \frac{\sum_0^4 xf(x)}{\sum_0^4 f(x)}$$

(c) Sea $p(x)$ la probabilidad de obtener el número x. Demuestre que x está dado por

$$(x)dx \sum_0^0 = \bar{x}$$

14. Considere la función

$$f(x) = \frac{1}{10}(10 - x)^2 \qquad 0 \leq x \leq 10$$

$$f(x) = 0 \qquad \text{para cualquier otra } x$$

(a) A partir de

$$\bar{x} = \frac{\int_{-\infty}^{\infty} xf(x)\,dx}{\int_{-\infty}^{\infty} f(x)\,dx}$$

encuentre el valor promedio de x. (b) Suponga que la variable x es discreta en lugar de continua. Además, $\Delta x = 1$, de modo que x sólo toma valores enteros 0, 1, 2,..., 10. Calcule $\bar{x}$ y compárelo con el resultado de la parte (a). (Sugerencia: puede ser más fácil calcular la suma apropiada directamente, en lugar de trabajar con fórmulas generales de sumas). (c) Calcule $\bar{x}$ para $\Delta x = 5$, es decir $x = 0, 5, 10$. Compárelo con el resultado de la parte (a). (b) Obtenga analogías entre los resultados obtenidos en este problema y el análisis de la sección 1-4. Asegúrese de entender los papeles que juegan $\bar{\mathscr{E}}$, $\Delta\mathscr{E}$, y $P(\mathscr{E})$.

15. Usando la relación $P(\mathscr{E}) = e^{-\mathscr{E}/kT}/kT$ y $\int_0^\infty P(\mathscr{E})\,d\mathscr{E} = 1$, evaluar la integral de (1-21), para deducir (1-22), $\bar{\mathscr{E}} = kT$.

16. Utilizando la relación $R_T(\nu)\,d\nu = (c/4)\rho_T(\nu)\,d\nu$ entre la radiancia espectral y la densidad de energía junto con la ley de radiación de Planck, derivar la ley de Stefan. Es decir, demuestre que:

$$R_T = \int_0^\infty \frac{2\pi h}{c^2} \frac{\nu^3\,d\nu}{e^{h\nu/kT} - 1} = \sigma T^4$$

donde $\sigma = 2\pi^5 k^4/15c^2h^3$

$$\left(\text{Sugerencia: } \int_0^\infty \frac{q^3\,dq}{e^q - 1} = \frac{\pi^4}{15}. \right)$$

17. Derive la ley del desplazamiento de Wien, $\lambda_{máx}\,T = 0.2014\,hc/k$, resolviendo la ecuación $d\rho(\lambda)/d\lambda = 0$. (Sugerencia: haga $hc/\lambda kT = x$ y demuestre que la ecuación citada conduce a $e^{-x} + x/5 = 1$. Después, demuestre que $x = 4.965$ es la solución).

18. Para verificar exprese mentalmente que la recientemente descubierta radiación universal de 3°K, se ajusta precisamente a un espectro de cuerpo negro, se decide medir $R_\nu(\lambda)$ desde una λ debajo de λ_{max}, donde su valor es $0.2R_\nu(\lambda_{max})$ hasta una λ por encima de λ_{max} donde su valor es otra vez $0.2R_\nu(\lambda_{max})$. ¿En qué intervalo de λ deberán hacerse las medidas?

2

Fotones-Propiedades corpusculares de la radiación

2.1 **Introducción** 47

Interacción de radiación con materia.

2.2 **El efecto fotoeléctrico** 47

Potencial de frenamiento; frecuencia de corte; ausencia de tiempo de retraso.

2.3 **Teoría cuántica de Einstein del efecto fotoeléctrico** 50

Fotones; cuantización de la energía del fotón, función de trabajo; reevaluación de la constante de Planck; espectro electromagnético; conservación del impulso.

2.4 **El efecto Compton** 55

Corrimiento Compton; derivación de las ecuaciones de Compton; longitud de onda de Compton; dispersión de Thomson; competencia entre las dispersiones Thomson y Compton.

2.5 **Naturaleza dual de la radiación electromagnética** 61

Difracción; desdoblamiento de personalidad de la radiación electromagnética; actitud contemporánea de los Físicos.

2.6 **Fotones y emisión de rayos X** 62

Emisión de rayos X; bremsstrahlung; relación de bremsstrahlung a efecto fotoeléctrico.

2.7 **Producción y aniquilación de pares** 65

Positrones; producción de pares electrón-positrón; aniquilación de pares; positronio; teoría de Dirac de los positrones.

2.8 Secciones transversales para absorción y dispersión de fotones 70

Definición de sección transversal; dependencia con la energía de las secciones transversales de la dispersión, fotoeléctrica, de producción de pares y total; atenuación expoencial; coeficientes y longitudes de atenuación.

Preguntas 74

Problemas 75

2

Fotones-Propiedades corpusculares de la radiación

2.1 Introducción

En este capítulo se examinarán los procesos en los cuales la radiación interacciona con la materia. Los procesos que implican la dispersión o absorción de radiación por materia son tres, a saber: el efecto fotoeléctrico, el efecto Compton y la producción de pares. Mientras que la producción de radiación se debe al bremsstrahlung o a la aniquilación de pares. En cada caso se obtendrá evidencia experimental de que la radiación es de carácter corpuscular en su interacción con la materia a diferencia de su naturaleza ondulatoria cuando se propaga. En el capítulo siguiente se estudiará una generalización de este resultado debida a de Broglie, y que conduce directamente a la mecánica cuántica. Parte del material de estos dos capítulos será una revisión de tópicos que el estudiante conoce de sus cursos de física elemental.

2.2 El efecto fotoeléctrico

Los experimentos que confirmaron por vez primera la existencia de las ondas electromagnéticas y la teoría electromagnética de Maxwell acerca de la propagación de la luz fueron realizados por Heinrich Hertz en 1886 y 1887. Es uno de aquellos hechos fascinantes y paradójicos en la historia de la ciencia es el que Hertz observó en el curso de sus experimentos, este efecto lo usaría Einstein después para contradecir otros aspectos de la teoría electromagnética clásica. Hertz descubrió que una descarga eléctrica entre dos electrodos ocurre más fácilmente cuando sobre uno de ellos incide luz ultravioleta. Lenard, continuando algunos experimentos de Hallwachs demostró en seguida que la luz ultravioleta facilita la descarga debido a que ocasiona la emisión de electrones desde la superficie del cátodo, a esta emisión de electrones desde una superficie por la acción de la luz se le denomina *efecto fotoeléctrico*.

En la figura 2-1 se muestra un aparato utilizado para estudiar el efecto fotoeléctrico. Una envoltura de vidrio encierra al aparato en un espacio al vacío, la luz incidente que es monocromática penetra a través de una ventana de cuarzo, incide sobre la placa de metal A y libera electrones, (llamados *fotoelectrones*). Los cuales si es que son atraídos a la copa de metal B mediante una diferencia de potencial V aplicada entre A y B se podrán detectar como una corriente. El amperímetro G sirve para medir esta corriente fotoeléctrica.

La curva a de la figura 2-2 es una gráfica de la corriente fotoeléctrica como función de la diferencia de potencial V para un aparato como el de la figura 2-1. Si V se hace lo suficientemente grande, la corriente fotoeléctrica alcanza cierto valor límite (valor de saturación), para el cual todos los fotoelectrones emitidos desde A son colectados por la copa B.

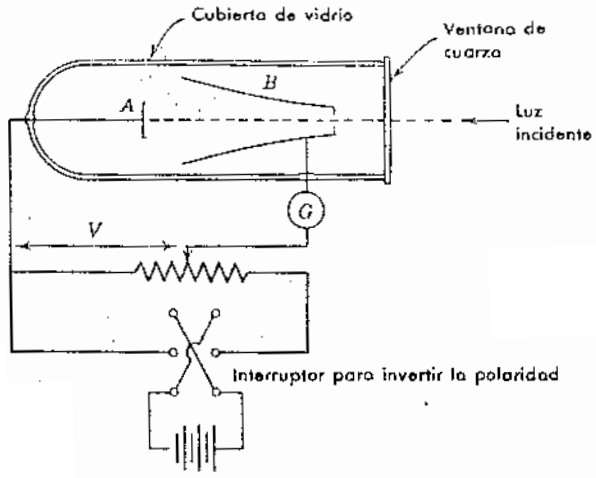


FIGURA 2-1

Aparato utilizado para estudiar el efecto fotoeléctrico. El voltaje V se puede variar en magnitud de manera continua y también es posible, mediante el interruptor, invertir el signo.

Si se invierte el signo de V, la corriente fotoeléctrica no cae inmediatamente a cero, lo cual sugiere que los electrones son emitidos desde A con energía cinética; algunos alcanzarán el punto B sin importar que el campo eléctrico se oponga a su movimiento. Sin embargo, si esta diferencia de potencial se hace lo suficientemente grande se alcanzará un valor V_0 para el cual la corriente fotoeléctrica se hace igual a cero, a este valor se le llama *potencial de frenamiento*. Esta diferencia de potencial V_0 multiplicada por la carga del electrón es una medida de la energía cinética $K_{máx}$ de los fotoelectrones *más rápidos*. Es decir

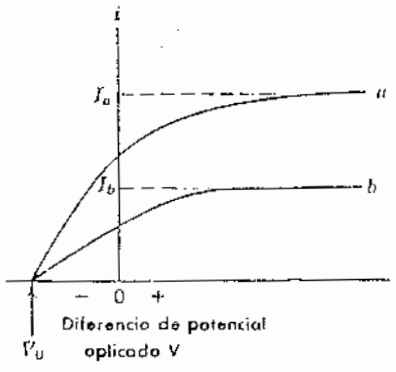
$$K_{máx} = eV_0 \tag{2-1}$$

Experimentalmente resulta que la cantidad $K_{máx}$ es independiente de la intensidad de la luz tal como la muestra la curva b de la figura 2-2 en la cual la intensidad de la luz se ha reducido a la mitad del valor utilizado para obtener la curva a.

La figura 2-3 muestra el potencial de frenamiento V_0 como función de la frecuencia de la luz que incide sobre el sodio. Obsérvese que existe una frecuencia de *corte* (ν_0) bien definida, por

FIGURA 2-2

Gráficas de la corriente I como función del voltaje V según los datos tomados con el aparato de la figura 2-1. La diferencia de potencial V se considera positiva cuando la copa B de la figura 2-1 es positiva con respecto a la superficie fotoeléctrica. En la curva b la intensidad de la luz incidente se ha reducido a la mitad de la de la curva a. El potencial de frenamiento V_0 es independiente de la intensidad de la luz pero las corrientes de saturación I_a e I_b son directamente proporcionales a él.



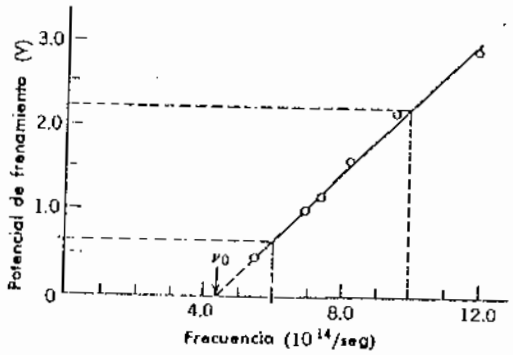


FIGURA 2-3

Gráfica de las medidas realizadas por Millikan para el potencial de frenamiento en función de la frecuencia para el caso del sodio. La frecuencia de corte ν_0 es 4.39×10^{14} Hz.

debajo de la cual no ocurre el efecto fotoeléctrico. Estos datos fueron tomados en 1914 por Millikan, quien por su esmerado trabajo sobre el efecto fotoeléctrico se hizo acreedor al premio Nobel de 1923. Debido a que el efecto fotoeléctrico para luz en el visible, o en el visible cercano es un fenómeno fundamentalmente de superficie, es necesario evitar en los experimentos capas de óxido, grasa o cualquier otro contaminante de la superficie.

Existen tres hechos fundamentales del efecto fotoeléctrico que no pueden explicarse en términos de la teoría clásica ondulatoria de la luz:

1. La teoría ondulatoria exige que el vector eléctrico oscilante E de la onda luminosa aumente en amplitud conforme aumenta la intensidad del haz luminoso. Puesto que la fuerza aplicada al electrón es eE, se sugiere que la energía cinética de los fotoelectrones también debería aumentar de acuerdo con el aumento en intensidad del haz luminoso. Sin embargo, la figura 2-2 muestra que K_{max}, que es igual a eV_0 *es independiente de la intensidad de la luz.* Esto ha sido probado sobre un intervalo de intensidades de 10^7.

2. Según la teoría ondulatoria, el efecto fotoeléctrico debería ocurrir para cualquier frecuencia de la luz tomando en cuenta solamente que la intensidad de la luz sea lo suficientemente intensa como para dar la energía necesaria para emitir los fotoelectrones. Sin embargo, la figura 2-3 muestra que para cada superficie existe una *frecuencia de corte característica* ν_0. *Para frecuencias menores de* ν_0, *el efecto fotoeléctrico no ocurre, sin importar qué tan intensa sea la iluminación.*

3. Si la energía adquirida por un fotoelectrón es absorbida de la onda que incide sobre la placa de metal, el "área efectiva del blanco" para un electrón en un metal está limitada y probablemente no sea mucho mayor que un círculo de un diámetro atómico. En la teoría clásica, la energía luminosa se encuentra uniformemente distribuida sobre el frente de onda. Entonces, si la luz es suficientemente débil existirá un tiempo de retraso mensurable entre el instante en que la luz empieza a incidir sobre la superficie y la expulsión del fotoelectrón, que se estimará en el ejemplo 2-1. Durante este intervalo el electrón deberá absorber energía del haz, hasta acumular la suficiente para escapar. *Sin embargo, nunca se ha medido un tiempo de retraso mensurable.* Este desacuerdo es particularmente interesante cuando la substancia fotoeléctrica es un gas, bajo estas circunstancias pueden aparecer mecanismos de absorción colectiva y la energía del fotoelectrón emitido deberá ser extraída del haz luminoso por un solo átomo o molécula.

Ejemplo 2-1. A una distancia de un metro de una fuente luminosa cuya potencia es 1 W = 1 joule/seg se coloca una placa de potasio. Supóngase que un fotoelectrón emitido puede recibir su energía de un área circular de la placa cuyo radio r es, por ejemplo, un radio atómico: $r \simeq 1 \times 10^{-10}$ m. La energía necesaria para extraer un electrón de la superficie del potasio es aproximadamente 2.1 eV = 3.4×10^{-19} joule (un electronvolt = 1 eV = 1.60×10^{-19} joules es la energía ganada por un electrón de carga $1.6 \times$

10^{-19} coulombs, que cae a través de una diferencia de potencial de 1 V). ¿Cuánta tiempo le tomaría a este blanco absorber esta cantidad de energía de la fuente luminosa? Supóngase que la energía luminosa está espareida uniformemente sobre el frente de onda.

El área del blanco es $\pi r^2 = \pi \times 10^{-20}$ m^2. El área de una esfera de un metro con centro en la fuente es $4\pi(1\text{ m})^2 = 4\pi$ m^2. Entonces, si la fuente radía unifarmemente en todas direcciones (es decir, si la energía está uniformemente distribuida sobre frentes de onda esféricos que emergen desde la fuente, según la teoría clásica) la razón R a la cual la energía cae sobre el blanco está dada par

$$R = 1 \text{ joule/seg} \times \frac{\pi \times 10^{-20}\text{ m}^2}{4\pi\text{ m}^2} = 2.5 \times 10^{-21} \text{ joule/seg}$$

Suponiendo que tada esta potencia ae absorbe, se puede calcular el tiempo necesaria para que los electranes adquieran energía suficiente para escapar; por lo tanto

$$t = \frac{3.4 \times 10^{-19}\text{ joule}}{2.5 \times 10^{-21}\text{ joule/seg}} = 1.4 \times 10^{2}\text{ seg} \simeq 2\text{ mia}$$

Por supuesto, se podría modificar la descripción anterior con objeta de reducir el tiempa calculada suponienda una mayor área efectiva del blanco. La supasición más favorable, que la energía se transfiere de la onda luminasa al electrón mediante un proceso resonante, canduce a una área del blanca de λ^2, dande λ es la longitud de anda de la luz, pero se podría abtener un tiempo de retrasa finita que estaría dentra dc nuestra habilidad experimental para medir, por ejempla $\lambda = 100$ Å, segundos. (Para luz ultravioleta de $t \simeq 10^{-2}$ seg Sin embarga, no se ha detectado ningún tiempo de retraso bajo ninguna circunstancia y los primeros experimentos fijaron un límite superiar de 10^{-9} segundos sobre cualquier retraso pasible. ◀

2.3 Teoría cuántica de Einstein del efecto fotoeléctrico

En 1905 Einstein puso en duda la teoría elásica de la luz, propuso uno nueva teoría y citó el efecto fotoeléctrico como una de las aplicaciones que podría probar cuál teoría era la correcta. Esto fue enunciado muchos años antes al trabajo de Millikan, pero Einstein estaba influenciado por el experimento de Leoard. Como ya se ha mencionado, ariginalmente Planck restringió su concepto de cuantización de la energía al electrón radiante en las paredes de una cavidad de cuerpo negro; Planek pensaba que la energía electromagnética una vez radiada se esparcía en el espacio de la misma manera que las ondas de agua se esparcen en ella. En vez de esto, Einstein propuso que la energía radiante estaba cuantizada en paquetes coocentrados a los cuales posteriormente se les llamó *fotones*.

Einstein argumentaba que los experimentos ópticos tan bien conocidos sobre interferencia y difracción de radiación electromagnética solamente se habían efectuado en situaciones en las cuales el número de fotones era muy grande. Estos experimentos dan resultados que son promedios de los comportamientos de los fatones individuales. En estos experimentos la presencia de los fotanes no se manifiesta, coma tampoco lo hacen las gotas de agua individuales en un rocío fino proveoiente de una manguera de jardío si el número de gotas es muy grande. Por supuesto que los experimentos sobre interferencia y difracción muestran de manera definitiva que los fotones no viajan desde donde son emitidos hasta donde son absorbidos en la forma simple en que lo hacen las partículas clásicas tales como las gotas de agua. Viajan como ondas clásicas en el sentido de que los cálculos que se basan en la forma en la que se propagan las andas (y en particular la forma en que dos ondas se refuerzan o nulifican entre sí dependiendo de sus fases relativas) explican correctamente las medidas del modo en el que viajan los fotones promedio.

Einstein no enfocó su atención en la forma ondulatoria de propagación de la radiación, sino que primero pensó en la forma en la cual se emiten y absorben los corpúsculos. Peosó que el requisito de Planck acerca de que la energía contenida en las ondas electromagnéticas de

frecuencia ν es una fuente radiante (por ejemplo una fuente de luz ultravioleta en un experimento fotocléctrico) sólo podía ser cero, o $h\nu$, o $2h\nu$, ... o $nh\nu$, ... implicoba que en el proceso de ir de un estado de energía $nh\nu$ a un estado de energía $(n-1)\,h\nu$ la fuente debería emitir pulsos de energía electromagnética discretos con contenida energético $h\nu$.

Einstein supuso que dicha paquete de energía está localizado inicialmente en un volumen de espacio pequeña y que se mantiene localizado mientras se mueve apartándose de la fuente con velocidad c. Supuso que la cantidad de energía E del puquete o fotón está relacionada con su frecuencia ν mediante la ecuución

$$E = h\nu \tag{2-2}$$

Tombién supuso que en el praceso fotoeléctrico un fotón es completamente absorbido por un electrón del fotocátodo.

Cuanda se emite un elcetrón desde la superficie del metal, su energía cinética es

$$K = h\nu - w \tag{2-3}$$

donde $h\nu$ es la energía del fotón incidente absorbido y w es el trabajo necesario para sacar al elecirón del metal. Este trabajo es el necesario para superar tanto los campos atractivas de los átomos en la superficie, cnmo las pérdidas de energía cinéticas debidas a las colisiones internas del electrón. Algunos de los electrones están ligados más fuertemente que atros y algunos pierden energía en colisiones durante el trayecto; en el easo en el que el enlace sea el más débil y no existan pérdidas internas, el fotoelectrón emergerá con la energía cinética máxima K_{max}. Por la tanto,

$$K_{max} = h\nu - w_0 \tag{2-4}$$

donde w_0 que es una energía característica del metal denaminada la *función trabajo* es la energía mínima necesnria para que el electrón pase a través de la superficie del metal y escape a las fuerzas atractivas que normalmente fijan el electrán al metal.

Ahora se contemplará la manera en la que la hipótesis de Einstein acerca del fotón encara las tres objeciones manifestadas en contra de la interpretación que la teoría ondulatoria hace del efecto fotocléctrico. Respecto a la primera objeción (la falta de dependencia de K_{max} con la intensidad de iluminación), existe completa concordancia entre la teoría del fotón y el experimento, puesto que duplicando la intensidad de la luz simplemente se duplica el número de fotones y por ende la corriente fotoeléctrica; esto no cambia la energía $h\nu$ de los fotones individuales o la naturaleza del proceso fotocléctrico individual descrito por (2-3).

Según la ecuación (2-4) la objeción 2 (la existencia de una frecuencia de corte) desaparece. Si K_{max} es igual a cero se tiene

$$h\nu_0 = w_0 \tag{2-5}$$

lo cual asevera que un fotón de frecuencia ν_0 tiene justamente la energía suficiente para extraer los fotoelectrones y nada extra que aparezca como energía cinética. Si la frecuencia se reduce por debajo de ν_0, los fotoelectrones individuales, sin importar cuántos haya, (esto es, sin importar la intensidad de la iluminación) no tendrán individualmente energía suficiente para extraer fotoelectrones.

En la teoría del fotón se hace a un lado la objeción 3 (la ausencia de un tiempo de retraso), debido a que la energía necesaria se suministra en paquetes concentrados y *no* se esparce uniformemente sobre un área grande como se supuso en el ejemplo 2-1, el cual está basado en la suposición de que la teoría ondulatoria clásica es cierta. Si existe cualquier iluminación incidiendo sobre el cátodo, entonces existirá cuanda menos un fotón que colinda con él y este fotón será absorbido inmediatamente por *algún* átomo resultando de inmediato en la emisián de un fotoelectrón.

Se volverá a escribir la ecuación de Einstein para el efecto fotoeléctrico, (2-4), substituyendo eV_0 por K_{max} de la ecuación (2-1). Se tiene

$$V_0 = \frac{h\nu}{e} - \frac{w_0}{e}$$

Por lo tanto la teoría de Einstein predice una relación lineal entre el potencial de frenamiento V_0 y la frecuencia ν en completa concordancia con los resultados experimentales que se muestran en la figura 2-3. La pendiente de la curva experimental en la figura debe ser h/e, o

$$\frac{h}{e} = \frac{2.20\text{ V} - 0.65\text{ V}}{10.0 \times 10^{14}/\text{seg} - 6.0 \times 10^{14}/\text{seg}} = 3.9 \times 10^{-15}\text{ V-seg}$$

Multiplicando esta relación por la carga electrónica e es posible determinar h, resultando $h = 3.9 \times 10^{-15}$ V-seg $\times$ 1.6×10^{-19} coul $= 6.2 \times 10^{-34}$ joule seg. De un análisis más cuidadoso y de estos y otros datos incluyendo algunos tomados con superficies de litio, Milliken encontró el valor $h = 6.57 \times 10^{-34}$ joules-seg con una precisión de aproximadamente 0.5%. Este resultado se ajusta con el valor de h derivado a partir de la fórmula de la radiación de Planck. El acuerdo numérico en dos determinaciones de h usando fenómenos y teorías completamente diferentes es sorprendente. Un valor de h obtenido recientemente partiendo de diversos experimentos es

$$h = 6.6262 \times 10^{-34}\text{ joule-seg}$$

Citando a Millikan: "El efecto fotoeléctrico... constituye una prueba que es totalmente independiente de los hechos de la radiación del cuerpo negro, de la validez de la suposición fundamental de la teoría cuántica, a saber la suposición de una emisión discontinua o explosiva de la energía absorbida por los constituyentes electrónicos de los átomos de las . . . ondas. Materializa, por así decirlo, la cantidad h descubierta por Planck mediante el estudio de la radiación del cuerpo negro y nos proporciona una confianza de que la concepción física fundamental que sustenta el trabajo de Planck corresponde a la realidad, confianza que no nos inspira ningún otro tipo de fenómeno".

Ejemplo 2-2. A partir de la figura 2-3 deducir la función trabajo del sodio.

La intersección de la recta con el eje horizontal es la frecuencia de corte $\nu_0 = 4.39 \times 10^{14}$/seg que substituido en (2-5) resulta en

$$w_0 = h\nu_0 = 6.63 \times 10^{-34}\text{ joule-seg} \times 4.39 \times 10^{14}/\text{seg}$$
$$= 2.92 \times 10^{-19}\text{ joule} \times \frac{1\text{ eV}}{1.60 \times 10^{-19}\text{ joule}}$$
$$= 1.82\text{ eV}$$

De la figura 2-3 se obtiene este mismo valor si se observa la intersección de la prolongación de la recta con el eje vertical. Sin embargo, los experimentos modernos dan un valor ligeramente mayor, 2.27 eV.

Para la mayoría de los metales conductores el valor de la función trabajo es del orden de unos cuantos electrónvolts; valor que es el mismo que el de la función trabajo para emisión termoiónica de estos metales. ◀

Ejemplo 2-3. ¿Cuál es la razón por unidad de área con la cual los electrones inciden sobre la placa de metal del ejemplo 2-1? Suponga que la luz es monocromática de longitud de onda 5890 Å (luz amarilla).

La razón por unidad de área con la cual caen los electrones sobre una placa de metal que dista 1 m de una fuente luminosa de un 1-W, es (ver ejemplo 2-1)

$$R = \frac{1\text{ joule/seg}}{4\pi(1\text{ m})^2} = 8.0 \times 10^{-2}\text{ joule/m}^2\text{-seg}$$
$$= 5.0 \times 10^{17}\text{ eV/m}^2\text{-seg}$$

Cada fotón tiene una energía

$$E = h\nu = \frac{hc}{\lambda} = \frac{6.63 \times 10^{-34}\ \text{joule-seg} \times 3.00 \times 10^{8}\ \text{m/seg}}{5.89 \times 10^{-7}\ \text{m}}$$

$$= 3.4 \times 10^{-19}\ \text{joule}$$

$$= 2.1\ \text{eV}$$

Por lo tanto, la razón R con la cual los fotones inciden sobre un área unitaria de la placa es

$$R = 5.0 \times 10^{17}\ \text{eV/m}^2\text{-seg} \times \frac{1\ \text{fotones}}{2.1\ \text{eV}} = 2.4 \times 10^{17}\ \frac{\text{fotones}}{\text{m}^2\text{-seg}}$$

La razón por la cual se presenta la posibilidad de que ocurra el efecto fotoeléctrico es que la energía del fotón es justamente 2.1 eV, que coincide con la función de trabajo para la superficie del potasio (ver ejemplo 2-1). Obsérvese, que si se incrementa ligeramente la longitud de onda (es decir si ν se disminuye ligeramente) no se presentará efecto fotoeléctrico, sin importar cuán grande sea la razón R.

Este ejemplo sugiere que la intensidad de la luz I se puede considerar como el producto de N, el número de fotones por unidad de área por unidad de tiempo, y $h\nu$, la energía de un solo fotón. Se ve que aún a intensidades relativamente bajas($\simeq 10^{-1}$ W/m^2)el número N es extremadamente grande ($\simeq 10^{17}$ fotones/m^2·seg) de manera que la energía para uno cualquiera de los fotones es muy pequeña. Esta da cuenta de la fineza extremada de la granularidad de la radiación y sugiere el por qué ordinariamente es tan difícil el detectarla. Esto es análogo a detectar la estructura atómica de la materia, la cual para la mayoría de los propósitos se puede considerar como continua; su carácter discreto se pone de manifiesto solamente bajo circunstancias especiales. ◀

En 1921 Einstein recibió el premio Nobel por la predicción teórica que hizo de la ley del efecto fotoeléctrico. Antes de que Millikan completara la confirmación experimental de esta ley (1914), Einstein fue recomendado por Planck y otros para ingresar como miembro en la Academia de Ciencias de Prusia. Su actitud negativa inicial acerca de la hipótesis del fotón se patentiza en su declaración firmada en elogio de Einstein, en la cual ellos escriben: "resumiendo, podemos decir que difícilmente existe entre los grandes problemas, en los que la física moderna es tan rica, uno al que Einstein no haya hecho una contribución importante. El que algunas veces haya perdido el objetivo en sus especulaciones, como por ejemplo en su hipótesis sobre los quanta de luz (fotones), realmente no dice mucho en su contra porque no es posible introducir ideas fundamentalmente nuevas aun en la más exacta de las ciencias sin correr ocasionalmente un riesgo".

En la actualidad la hipótesis del fotón se usa en todo el espectro electromagnético y no sólo en la región luminosa (véase figura 2-4). Por ejemplo, se puede decir que una cavidad de microondas contiene fotones. Para $\lambda = 10$ cm, una longitud de onda típica de microondas, se puede calcular que la energía del fotón es 1.20×10^{-5} eV. Esta energía es demasiado pequeña para extraer fotoelectrones de superficies metálicas. Para rayos X o para rayos λ energéticos tales como los emitidos por núcleos radiactivos la energía del fotón puede ser de 10^6 eV o más. Esos fotones pueden extraer electrones internos ligados a átomos pesados con energías del orden de 10^5 eV. Los fotones en la región visible del espectro electromagnético no son lo suficientemente energéticos como para hacer esto, los fotoelectrones que ellos extraen provienen de los llamados electrones de *conducción* los cuales están ligados al metal con energías de unos cuantos electrónvolts.

Obsérvese que en el proceso fotoeléctrico los fotones son absorbidos, lo cual exige que los electrones estén ligados a los átomos o en sólidos, ya que un electrón realmente libre no puede absorber un fotón y conservar en el proceso, tanto su energía relativista total como su impulso. Por lo tanto, se deberá tener un electrón ligado de manera que las fuerzas de ligadura sirvan para transmitir el impulso al átomo o al sólido. Debido a la gran masa de un átomo o un sólido, comparada con la del electrón, el sistema puede absorber una gran cantidad de impulso sin adquirir una apreciable cantidad de energía. La ecuación de energía en el efecto fotoeléctrico

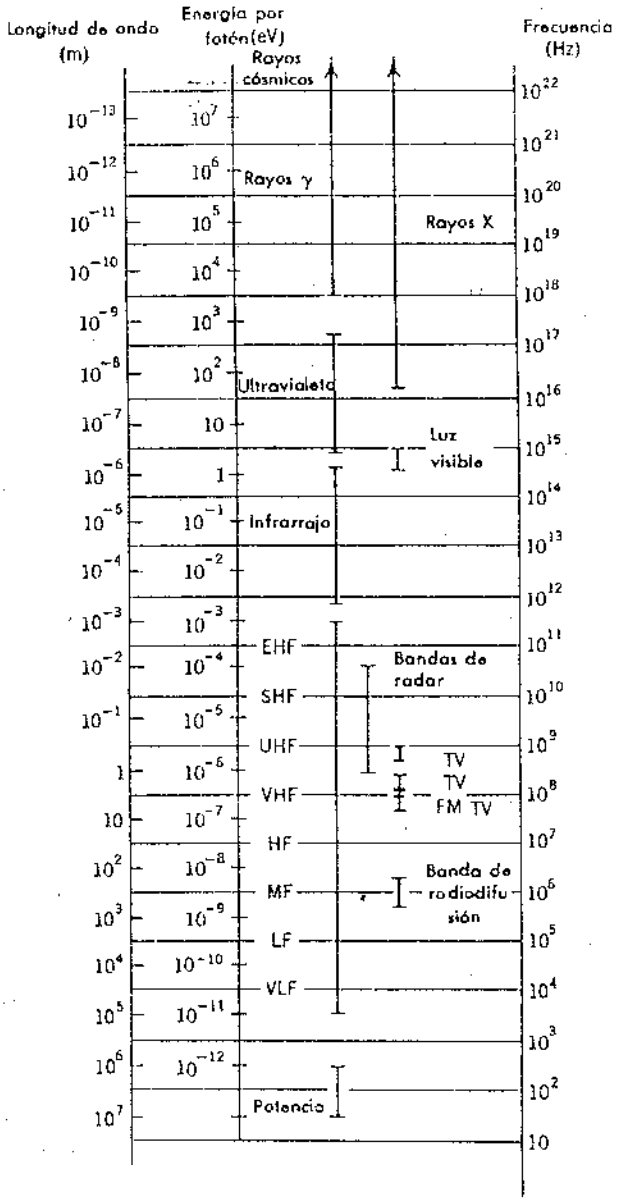


FIGURA 2-4

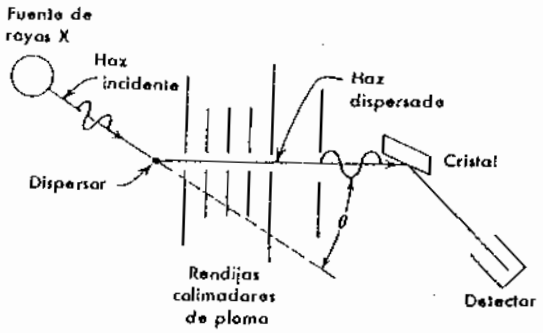
El espectro electromagnético mostrando la longitud de onda, la frecuencia y la energía por fotón sobre una escala logarítmica.

sigue siendo válida, el efecto sigue siendo posible sólo porque existe una partícula pesada que retrocede además de un electrón emitido. El efecto fotoeléctrico es una forma importante en la cual los fotones, de energías comparables con las de los rayos X o superiores, son absorbidas por la materia. A energías mayores resultan más importantes otros procesos de absorción de fotón tal como se verá posteriormente.

Por último, se debe hacer énfasis que en el modelo de Einstein un fotón de frecuencia ν tiene exactamente la energía $h\nu$ y *no* tiene energías que sean múltiplos enteros de $h\nu$. Por supuesto que puede haber n fotones de frecuencia ν de tal manera que la energía a tal frecuencia puede

FIGURA 2-5

Arreglo experimental utilizado por Compton. Los rayos X monocromáticos de longitud de onda λ inciden sobre un dispersor de grafito. La distribución de intensidades respecto a la longitud de onda se mide mediante los rayos X dispersados para cualquier ángulo de dispersión θ. Las longitudes de onda dispersadas se miden observando las reflexiones de Bragg del cristal (ver fig. 3-3). Sus intensidades se miden mediante un detector tal como una cámara de ionización.



ser $nh\nu$. En el tratamiento de la radiación de la cavidad de cuerpo negro según el modelo de Einstein se trabaja con un "gas de fotones". Debido a que la energía radiante está localizada en el espacio en paquetes en vez de estar extendida a través del espacio en ondas estacionarias. Años después de la deducción de Planck de la fórmula para la radiación de la cavidad, Bose y Einstein derivaron la misma fórmula trabajando con un gas de fotones.

2.4 El efecto Compton

En el año de 1923 la naturaleza corpuscular de la radiación recibió una confirmación dramática proveniente de los experimentos de Compton. Como se muestra en la figura 2-5, Campton permitía que un haz de rayos X de longitud de anda λ perfectamente definida incidiera sobre un blanco de grafito. Para varios ángulos de disperaión, medía la intenaidad de los rayas X dispersados como una función de su longitud de onda y sus resultadaa experimentales ae muestran en la figura 2-6. Se observa que no obstante que el haz incidente consiste esencialmente de una sala longitud de onda λ, loa rayos X dispersados tienen picos de intensidad para *dos* langitudes de onda; una de las cuales ea la misma longitud de onda incidente y la otra, λ', es mayor por una cantidad $\Delta\lambda$. Este corrimiento, llamado *corrimiento Compton*, $\Delta\lambda = \lambda' - \lambda$ varía can el ángula al que se observan las rayos X disperaados.

La presencia de una longitud de onda dispersada λ' no es pasible entenderla si se considera a la radiación x incidente coma una onda electromagnétiea clásica. En el modelo clásico, el vector ascilante de campa eléctrica en la onda incidente de frecuencia ν actúa sobre los electrones libres del blanco y los pone a oscilar a la misma frecuencia. Estos electrones oscilantes, como las cargas que van y vienen en una pequeña antena transmisora de radia, radían ondas electromagnéticas que tienen nuevamente esta misma frecuencia ν. Por lo tanto, en el modelo clásica la anda dispersada deberá tener la misma frecuencia ν y la misma longitud de onda λ que la onda incidente.

Compton (y Debye en farma independiente) interpretó su resultado experimental postulando que el haz de rayas X incidente no era una onda de frecueneia ν sino una colección de fotones cada una de los cuales con energía $E = h\nu$ y que estos fotones chocaban con los electrones libres del blanco dispersor de mancra similar a la colisión entre bolas de billar. Según este punto de vista, los fotones de "retroceso" que emergen del blanco constituyen la radiación dispersada. Puesto que el fotón incidente transfiere algo de su energía al electrón con el cual choca, el fotán dispersado deberá tener una energía menor E'; y par lo tanto deberá tener una frecuencia menor $\nu' = E'/h$, lo cual implica una longitud de onda mayar $\lambda' = c/\nu'$. Esto viene a dar cuenta de manera cualitativa del corrimiento en langitud de anda $\Delta\lambda = \lambda' - \lambda$. Obsérvese que en la interacción se considera a los rayus X enmo partículas y no como ondas y que a diferencia de su comportamiento en el proceso fotoeléctrico los fotones de rayos X son

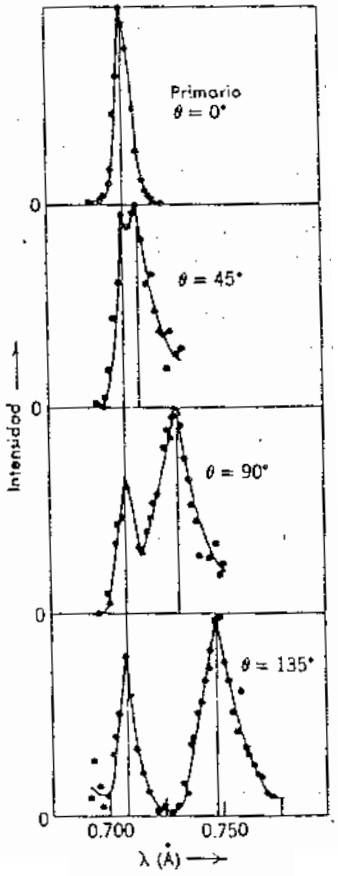


FIGURA 2-6

Resultados experimentales de Compton. La línea continua vertical de la izquierda corresponde a la longitud de onda λ, y la de la derecha a λ'. Se muestran los resultados para cuatro distintos ángulos de dispersión (θ). Obsérvese que el corrimiento Compton $\Delta\lambda = \lambda' - \lambda$, para $\theta = 90°$, concuerda con la predicción teórica, 0.0243 Å.

dispersados en vez de ser absorbidos. Analicemos ahora cuantitativamente una sola colisión fotón-electrón.

Para una radiación x de frecuencia ν la energía de un fotón en el haz incidente es

$$E = h\nu$$

Tomando literalmente la idea de un fotón como un paquete de energía localizado, habrá que considerarlo como una partícula de energía E e impulso p; sin embargo, tal partícula deberá tener propiedades muy especiales. Considérese la ecuación (ver apéndice A) que da la energía relativista total de una partícula en términos de su masa en reposo m_0 y su velocidad v.

$$E = m_0c^2/\sqrt{1 - v^2/c^2}$$

Puesto que la velocidad de un fotón es igual a c y puesto que su contenido energético $E = h\nu$ es finito, resulta evidente que la masa en reposo de un fotón deberá ser cero. Por lo tanto, se puede considerar al fotón como una partícula de masa en reposo cero y de energía total relativista E la cual es por completo energía cinética. El impulso de un fotón se puede evaluar de la relación general entre la energía relativista total E, el impulso p y la masa en reposo m_0. Es decir,

$$E^2 = c^2p^2 + (m_0c^2)^2 \tag{2-6}$$

Para un fotón, el segundo término del segundo miembro es igual a cero y entonces se tiene

$$p = E/c = h\nu/c \tag{2-7}$$

o

$$p = h/\lambda \tag{2-8}$$

donde $\lambda = c/\nu$ es la longitud de onda de la radiación electromagnética que constituye al fotón. Resulta interesante observar que la teoría ondulatoria clásica de Maxwell para la radiación electromagnética conduce también a una ecuación $p = E/c$, en la que p representa el impulso de la radiación por unidad de volumen y E representa su energía por unidad de volumen.

Se observa que la frecuencia ν de la radiación dispersa es independiente del material blanco, lo cual implica que la dispersión no involucra átomos completos. Compton supuso que la dispersión era debida a las colisiones entre el fotón y un electrón individual en el blanco, también supuso que los electrones que participan en este proceso de dispersión son electrones libres e inicialmente estacionarios. Se pueden encontrar justificaciones *a priori* a estas suposiciones considerando el hecho de que la energía de un fotón de rayos X es mayor que la energía de un fotón ultravioleta en varios órdenes de magnitud y de la discusión del efecto fotoeléctrico resulta evidente que la energía de un fotón ultravioleta es comparable a la energía mínima con la cual un electrón está ligado en un metal.

Considérese, como en la figura 2-7, una colisión entre un fotón y un electrón libre estacionario. En el diagrama de la izquierda se muestra un fotón de energía relativista E_0 e impulso p_0 que incide sobre un electrón estacionario de energía en reposo m_0c^2. En el diagrama de la derecha el fotón ha sido dispersado a un ángulo θ y se mueve con energía relativista total E_1 e impulso p_1 mientras que el electrón retrocede a un ángulo φ con energía cinética K e impulso p. Compton aplicó la conservación del impulso y de la energía relativista total a este problema de colisiones. Y utilizó ecuaciones relativistas ya que el fotón siempre se mueve con velocidades relativistas y el electrón que retrocede también se mueve, en la mayoría de los casos, bajo estas circunstancias.

La conservación del impulso exige

$$p_0 = p_1 \cos\theta + p\cos\varphi$$

$$p_1 \operatorname{sen}\theta = p \operatorname{sen}\varphi$$

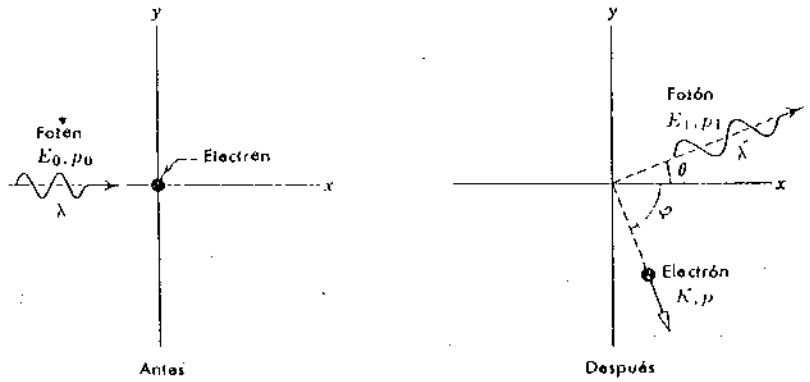


FIGURA 2-7

Interpretación de Compton. Un fotón de longitud de onda λ incide sobre un electrón libre en reposo. En la colisión el fotón es dispersado a un ángulo θ con una longitud de onda mayor λ' mientras que el electrón emerge a un ángulo φ.

Elevando al cuadrado estas dos ecuaciones se obtiene

$$(p_0 - p_1 \cos \theta)^2 = p^2 \cos^2 \varphi$$

y

$$p_1^2 \operatorname{sen}^2 \theta = p^2 \operatorname{sen}^2 \varphi$$

y sumando

$$p_0^2 + p_1^2 - 2p_0 p_1 \cos \theta = p^2 \tag{2-9}$$

La conservación de la energía relativista total requiere

$$E_0 + m_0 c^2 = E_1 + K + m_0 c$$

entonces

$$E_0 - E_1 = K$$

Y de acuerdo con (2-7), se tiene

$$c(p_0 - p_1) = K \tag{2-10}$$

Si en la ecuación (2-6) se escribe $K + m_0 c^2$ en vez de E se obtiene

$$(K + m_0 c^2)^2 = c^2 p^2 + (m_0 c^2)^2$$

Que se simplifica a

$$K^2 + 2K m_0 c^2 = c^2 p^2$$

o

$$K^2/c^2 + 2K m_0 = p^2$$

Evaluando p^2 de (2-9) y K de (2-10) se obtiene

$$(p_0 - p_1)^2 + 2m_0 c(p_0 - p_1) = p_0^2 + p_1^2 - 2p_0 p_1 \cos \theta$$

que se reduce a

$$m_0 c(p_0 - p_1) = p_0 p_1 (1 - \cos \theta)$$

o

$$\frac{1}{p_1} - \frac{1}{p_0} = \frac{1}{m_0 c}(1 - \cos \theta)$$

Multiplicando por h y utilizando (2-8), se obtiene la *ecuación de Compton*

$$\Delta\lambda = \lambda_1 - \lambda_0 = \lambda_C(1 - \cos \theta) \tag{2-11}$$

donde

$$\lambda_C \equiv h/m_0 c = 2.43 \times 10^{-12} \text{ m} = 0.0243 \text{ Å} \tag{2-12}$$

es la llamada *longitud de onda de Compton*.

Obsérvese que el *corrimiento Compton* $\Delta\lambda$, sólo depende del ángulo de dispersión θ y *no* de la longitud de onda inicial λ. La ecuación (2-11) predice los corrimientos Compton de la figura 2-6 observados experimentalmente dentro de los límites de la precisión experimental. Según (2-

11) $\Delta\lambda$ varía desde cero (para $\theta = 0$ que corresponde a una colisión "rasante" en la que el fotón incidente apenas si se deflecta) hasta $2h/m_0c = 0.049$ Å (para $\theta = 180°$ que corresponde a una colisión de "frente" en la que el fotón incidente regresa por la misma dirección por donde vino). La figura 2-8 es una gráfica de $\Delta\lambda$ contra θ.

En experimentos subsecuentes (realizados por Compton, Simon, Wilson, Bothe, Geiger y Blass) se detectó el electrón de retroceso resultante, mostrando que éste aparece simultáneamente con los rayos X dispersadas y confirmando cuantitativamente la predicción de la energía y dirección del electrón dispersa.

Aún deberá explicarse la presencia de un pico en la figura 2-6 para el cual la longitud de onda del fotón no cambia en la dispersión. Hasta aquí se ha supuesto que el electrón con el cual el fotón choca es un electrón libre. Esta suposición es justificable aun cuando el electrón esté ligado inicialmente, si la energía cinética que adquiere éste durante la colisión es mucho mayor que su energía de enlace. Si el electrón está fuertemente ligado a un átomo del blanco o si la energía del fotón incidente es muy pequeña, existe alguna posibilidad de que el electrón no sea expulsado del átomo; en tal caso se puede considerar que la colisión tiene lugar entre el fotón y el átoma completo. La coraza iónica a la cual está ligado el electrón en el dispersor blanco retrocede como un todo durante la colisión, entonces la masa M del átomo es la masa característica del proceso y deberá substituirse por la masa del electrón m_0 en las ecuaciones de corrimiento de Compton. Puesto que $M \gg m_0$ ($M \simeq 22{,}000\ m_0$ para el carbón, por ejemplo) el corrimiento Compton para colisiones con electrones fuertemente ligados es, como se puede ver de (2-11) y (2-12) inconmensurablemente pequeño (una millonésima de Angstrom para el carbón) de tal manera que el fotón dispersado no modifica su longitud de onda esencialmente. En resumen, algunos fotones son dispersados por electrones que a su vez son liberados por la colisión; estos fotones modifican su longitud de onda, por otra parte, hay otros fotones que son dispersados por electrones que permanecen ligados durante la colisión; en estos fotones no se modifica su longitud de onda.

Al proceso de fotones dispersados sin cambio en su longitud de onda se le llama *dispersión Thomson* en honor del físico que desarrolló la teoría clásica de la dispersión de rayos X por átomos, por el año 1900. Thomson consideró los rayos X como un haz de ondas electromagnéticas cuyo campo eléctrico oscilante interactuaba con las cargas de los electrones atómicos del blanco. Esta interacción produce fuerzas sobre los electrones que ocasionan aceleraciones

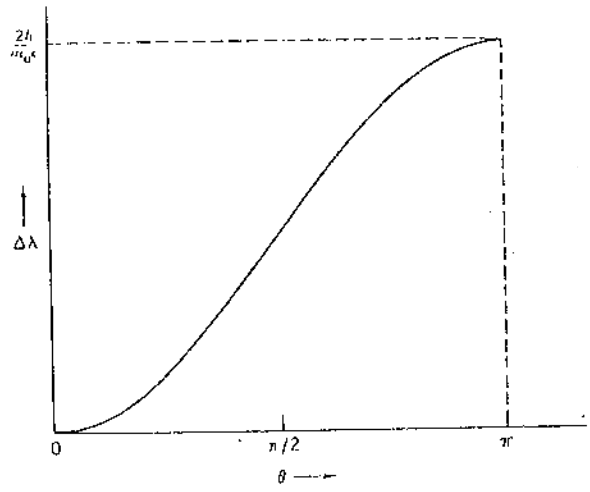


FIGURA 2-8

Resultado de Compton. $\Delta\lambda = (h/m_0c)(1 - \cos\theta)$.

oscilantes; como resultado de las aceleraciones, los electrones radiarán ondas electromagnéticas de la misma frecuencia y en fase con las ondas incidentes, por lo tanto los electrones atómicos absorben energía del haz incidente de rayos X y lo dispersan en todas direcciones sin modificar la longitud de onda. No obstante que esta explicación clásica de la dispersión es diferente a la explicación cuántica presentada en el párrafo anterior, ambas explican los mismos hechos observados en las mediciones. Por lo tanto, la dispersión de Thomsan es un caso en el que concuerdan los resultados clásico y cuántico.

Es interesante el preguntar en qué región del espectro electromagnético será dominante el proceso de dispersión Thomson y en cuál el proceso Compton. Si la radiación incidente está en el espectro visible, microondas o radio, λ es extremadamente grande comparada con el corrimiento Compton $\Delta\lambda$, e independientemente, según la ecuación (2-12), de si se usa un electrón o una masa atómica para evaluar la longitud de onda Compton. Por consiguiente, la radiación dispersada en esta región del espectro tendrá en todas las circunstancias una longitud de onda que es la misma que la longitud de onda de la radiación incidente dentro de la precisión experimental. Así, cuando $\lambda \to \infty$ los resultados cuántico y clásico coinciden y la dispersión Thomson es la dominante. En la región del espectro que corresponde a los rayos X la dispersión Compton empieza a ser importante particularmente para blancos dispersores de número atómico bajo donde los electrones atómicos no están muy fuertemente ligados y el corrimiento de longitud de onda en la dispersión por un electrón que es liberado en el proceso empieza a ser fácilmente mensurable. En la región de rayos γ en la que $\lambda \to 0$, la energía del fotón se hace tan grande, que durante la colisión siempre hay un electrón liberado y domina la dispersión Compton.

Los resultados clásicos fallan al explicar la dispersión de radiación en la región de longitudes de onda cortas, como en la catástrofe ultravioleta de la física clásica donde las predicciones acerca de la radiación en una cavidad divergen radicalmente con los resultados experimentales para longitudes de onda cortas. Esto es debido al tamaño de la constante de Planck h. Para longitudes de onda largas la frecuencia ν es muy pequeña y como h también es pequeña la granularidad en la energía electromagnética, $h\nu$, es tan pequeña como para que sea virtualmente indistinguible del continuo de la física clásica. Pero a longitudes de onda suficientemente cortas donde ν es lo suficientemente grande, h ya no es tan pequeña como para ignorar los efectos cuánticos.

Ejemplo 2-4. Considérese un haz de rayos X con $\lambda = 1.000$ Å y también un haz de rayos γ provenientes de una muestra Cs^{137} con $\lambda = 1.88 \times 10^{-2}$ Å. Si la radiación dispersada por los electrones libres se observa a 90° del haz incidente: (a) ¿Cuál es el corrimiento en longitud de onda Compton en cada caso? (b) ¿Qué energía cinética se le comunica al electrón de retroceso en cada caso? (c) ¿Qué porcentaje de la energía del fotón incidente se pierde en la colisión en cada caso?

(a) El corrimiento Compton con $\theta = 90°$, es

$$\Delta\lambda = \frac{h}{m_0 c}(1 - \cos\theta) = \frac{6.63 \times 10^{-34}\ \text{joule-seg}}{9.11 \times 10^{-31}\ \text{kg} \times 3.00 \times 10^{8}\ \text{m/seg}} \times (1 - \cos 90°)$$
$$= 2.43 \times 10^{-12}\ \text{m} = 0.0243\ \text{Å}$$

Este resultado es independiente de la longitud de onda incidente y por lo tanto el mismo para rayos γ y para rayos X.

(b) La ecuación (2-10) se puede escribir como

$$hc/\lambda = hc/\lambda' + K$$

Entonces, puesto que $\lambda' = \lambda + \Delta\lambda$, se tiene

$$hc/\lambda = hc/(\lambda + \Delta\lambda) + K$$

de modo que $K = hc\,\Delta\lambda/\lambda(\lambda + \Delta\lambda)$.

Para el haz de rayos X con $\lambda = 1.00$ Å, se tiene

$$K = \frac{6.63 \times 10^{-34}\,\text{joule-seg} \times 3.00 \times 10^{8}\,\text{m/seg} \times 2.43 \times 10^{-12}\,\text{m}}{1.00 \times 10^{-10}\,\text{m} \times (1.00 + 0.024) \times 10^{-10}\,\text{m}} = 4.73 \times 10^{-17}\,\text{joule}$$
$$= 295\,\text{eV} = 0.295\,\text{keV}$$

Para el haz de rayos γ con $\lambda = 1.88 \times 10^{-2}$ Å, se tiene

$$K = \frac{6.63 \times 10^{-34}\,\text{joule-seg} \times 3.00 \times 10^{8}\,\text{m/seg} \times 2.43 \times 10^{-12}\,\text{m}}{1.88 \times 10^{-12}\,\text{m} \times (0.0188 + 0.0243) \times 10^{-10}\,\text{m}} = 5.98 \times 10^{-14}\,\text{joule}$$
$$= 378\,\text{keV}.$$

(c) La energía del fotón de rayos X incidente es

$$E = h\nu = \frac{hc}{\lambda} = \frac{6.63 \times 10^{-34}\,\text{joule-seg} \times 3.00 \times 10^{8}\,\text{m/seg}}{1.00 \times 10^{-10}\,\text{m}} = 1.99 \times 10^{-15}\,\text{joule}$$
$$= 12.4\,\text{keV}$$

La energía perdida por el fotón es igual a la energía ganada por el electrón, 0.295 keV, de modo que la pérdida porcentual en energía es

$$\frac{0.295\,\text{keV}}{12.4\,\text{keV}} \times 100\% = 2.4\%$$

La energía del fotón del rayo γ incidente es

$$E = h\nu = \frac{hc}{\lambda} = \frac{6.63 \times 10^{-34}\,\text{joule-seg} \times 3.00 \times 10^{8}\,\text{m/seg}}{1.88 \times 10^{-12}\,\text{m}} = 1.06 \times 10^{-13}\,\text{joule}$$
$$= 660\,\text{keV}$$

La energía perdida por el fotón es igual a la ganada por el electrón, 378 keV, de onda que la pérdida porcentual en energía es

$$\frac{378\,\text{keV}}{660\,\text{keV}} \times 100\% = 57\%$$

En consecuencia, los fotones con más contenido energético (los cuales tienen longitudes de onda pequeñas) experimentan una mayor pérdida *porcentual* de energía en la dispersión Compton. Esto corresponde al hecho de que los fotones de longitud de onda menor experimentan un incremento porcentual mayor en su longitud de onda al ser dispersados. Esto resulta evidente de la expresión para la pérdida fraccional en energía, dada simplemente por

$$\frac{K}{E} = \frac{hc\,\Delta\lambda/\lambda(\lambda + \Delta\lambda)}{hc/\lambda} = \frac{\Delta\lambda}{\lambda + \Delta\lambda}$$

De esto se puede demostrar que para $\lambda = 5500$ Å, que corresponde a fotones visibles, la pérdida porcentual (para $\theta = 90°$) es menor que un centésimo del 1% mientras que para $\lambda = 1.25 \times 10^{-2}$ Å, que corresponde a fotones de rayos γ de 1 MeV la pérdida porcentual (para $\theta = 90°$) es 67%. ◀

2.5 La naturaleza dual de la radiación electromagnética

En su artículo "teoría cuántica de la dispersión de rayos X por elementos ligeros" Compton escribió: "La presente teoría depende esencialmente de la suposición de que cada electrón que

es efectivo en la dispersión, dispersa un quantum completo (fotón). También involuera la hipótesis de que un quantum de radiación se recibe desde direeciones definidas y se dispersa en direeciones definidas. El soporte experimental de la teoría indica de manera muy convincente que un quantum de radiación lleva consigo su impulso y su energía".

Es olvia la necesidad de interpretación en procesos que tratan de la interacción entre la radiación y la materia para un fotón o partícula localizada; pero al mismo tiempo es necesario una teoría ondulatoria de la radiación para entender los fenómenos de interferencia y difracción. Por consiguiente debe tamarse en serio la idea de que la radiación no es un fenómeno puramente ondulatoric ni es nada más un chorro de partículas. Sea lo que fuere la radiación, ésta se comporta como ondulatoria bajo algunas circunstancias y como corpuscular bajo otras. Efectivamente, esta situación se revela muy claramente en el trabajo experimental de Comptan en el que; (a) se usa un espectrómetro de cristal para medir las longitudes de onda de los rayos X y las medidas se interpratan mediante una teoría ondulatoria de la difracción y (b) las dispersiones afectan la longitud de onda de una manera que sólo puede entenderse si se trata a los rayos X como partículas. Es en las expresiones $E=h\nu$ y $p = h/\lambda$ donde se combinan los atributos de las ondas (ν y λ) y los atributos de las partículas (E y p).

Sin embargo, muchos físicos se sintieron muy incómodos cuando contemplaron el "desdoblamiento de personalidad" de la radiación electromagnética, el punto de vista amplio que se ha proporcionadn por el desarrnllo de la mecánica cuántica ha ocasionado que la actitud contemporánea sea muy diferente. La evidente dualidad en la naturaleza onda-partícula de la radiación ya nn se considera tan desusual debido a que ahora se conoce que es una caraeterística general de todos los entes físicos. Se verá por ejemplo que los electrones y los protones tienen exactamente la misma naturaleza dual que los fntones. También se verá que es posible reconciliar la existencia de aspecios ondulatorios con la existencia de aspectos corpusculares para cualesquiera de estos entes con la ayuda de la mecánica cuántica.

2.6 Fotones y emisión de rayos X

Los rayos X, llamados así por su descubridor Roentgen debido a que entonces su naturaleza era desconocida, son radiaciones en el espectrn electrnmagnético de longitudes de onda algo menores que 1.0 Å. Muestran el comportamiento típico de una onda transversal, como polarización, interferencia y difracción que se encuentran en la luz y en todas las otras radiaciones electromagnéticas. Los rayos X se producen en el blanco de un tubo de rayos X como el mostrado en la figura 2-9, cuando un haz de electrones de alta energía, acelerados a través de una diferencia de potencial de miles de volts se frenan al chocar con el blanco. Según la física clásica (ver apéndice B) la desaceleración de los electrones hasta alcanzar el punto de

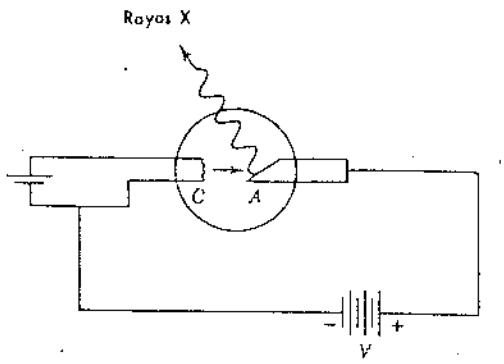


FIGURA 2-9

Tubo de rayos X. Los electrones son emitidos térmicamente desde el cátodo calentado C y se aceleran hacia el ánodo A mediante el potencial aplicado V. Los rayos X se emiten desde A cuando los electrones son detenidos por él.

reposo en el material blanco resulta en la emisión de un espectro contínuo de radiación electromagnética.

La figura 2-10 muestra, para cuatro valores distintos de la energía del electrón incidente, como se distribuyen en longitud de onda los rayos X emitidos desde el blanco de tungsteno. (Además del espectro continuo de rayos X también son emitidas líneas de rayos X características del material blanco. Estas líneas serán analizadas en el capítulo 9). La característica más notable de estas curvas es que para una energía determinada del electrón existe una longitud de onda mínima bien definida λ_{min}; por ejemplo, para electrones de 40 keV, λ_{min} es 0.311 Å. No obstante que la forma de la distribución del espectro continuo de rayos X depende ligeramente del material blanco escogido así como del potencial V que acelera los electrones, el valor de λ_{min} depende sólo de V y es el mismo para todos los materiales blancos. La teoría electromagnética clásica no puede dar cuenta de este hecho ya que no existe razón por la cual ondas cuya longitud de onda es menor que cierto valor crítico no deban salir del blanco.

Sin embargo, existe una explicación si se considera a los rayos X como fotones. La figura 2-11 muestra el proceso elemental que desde el punto de vista del fotón, es responsable del espectro continuo de rayos X que muestra la figura 2-10. Un electrón de energía cinética inicial K es desacelerado durante una colisión con un núcleo pesado, la energía que éste pierde aparece en forma de radiación como un fotón de rayos X. El electrón interactúa con el núclea cargado mediante el campo de Coulomb y transfiere impulso al núcleo. La desaceleración del electrón produce la emisión del fotón. El núcleo blanco es tan masivo, que la energía que adquiere durante la colisión se puede ignorar sin ningún problema. Si K' es la energía cinética del electrón después del choque, la energía del fotón es entonces

$$h\nu = K - K'$$

y la longitud de onda para el fotón se sigue de

$$hc/\lambda = K - K' \qquad (2\text{-}13$$

Los electrones del haz incidente pueden perder diferentes cantidades de energía en las colisiones y un determinado electrón podrá alcanzar el punto de reposo solamente después de muchas de ellas. Entonces los rayos X producidos por muchos electrones son los que constituyen el espectro continuo de la figura 2-10 y son muchos los fotones discretos cuyas

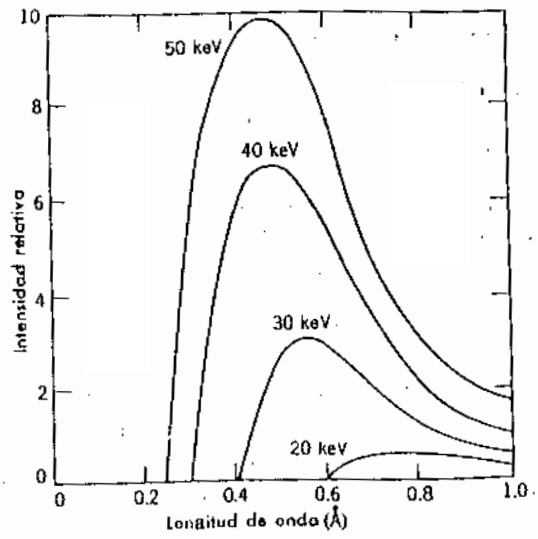


FIGURA 2-10

El espectro continuo de rayos X emitidos por un blanco de tungsteno para cuatro diferentes valores de la energía del electrón incidente (eV).

longitudes de onda varían desde λ_{min} hasta $\lambda \to \infty$, correspondiendo a las diferentes pérdidas de energía en los choques individuales. El fotón de longitud de onda más corta debe emitirse cuando un electrón pierde toda su energía cinética en un proceso de desaceleración; entonces $K' = 0$ de modo que $K = hc/\lambda_{min}$. Ya que K es igual a eV la energía adquirida por el electrón al acelerarse mediante la diferencia de potencial V aplicada al tubo de rayos X, se tiene

$$eV = hc/\lambda_{min}$$

o

$$\lambda_{min} = hc/eV \tag{2-14}$$

Entonces el mínimo de la longitud de onda de corte representa la conversión completa de la energía cinética del electrón en radiación X. La ecuación 2-14 muestra claramente que si $h \to 0$ entonces $\lambda_{min} \to 0$, la cual es la predicción de la teoría clásica. Esto muestra que la existencia de una longitud de onda mínima es un fenómeno cuántico.

Al espectro continuo de radiación X de la figura 2-10 se le llama frecuentemente *bremsstrahlung*, del alemán *bremss* (igual a frenar, es decir, desacelerar) más *strahlung* (igual a radiación). El proceso de bremsstrahlung ocurre no solamente en tubos de rayos X sino en cualquier parte en que los electrones choquen con la materia, como en los rayos cósmicos, en las cinturones de radiación de van Allen que circundan la tierra y en el frenamiento de electrones que emergen de los aceleradores o núcleos radiactivos. El proceso de bremsstrahlung se puede considerar como un efecto fotoeléctrico inverso: en el efecto fotoeléctrico, se absorbe un fotón cuya energía e impulso van al electrón y al núcleo que retrocede; en el proceso de bremsstrahlung se crea un fotón cuya energía e impulso vienen de la colisión de un electrón y un núcleo. En el proceso bremsstrahlung se trata de la *creación* de fotones en vez de su absorción o dispersión por la materia.

Ejemplo 2-5. Determinar la constante de Planck h a partir del hecho de que la longitud de onda mínima de rayos X producida por electrones de 40 keV es 3.11×10^{-11} m.

De 2.14, se tiene

$$h = \frac{eV\lambda_{min}}{c}$$

$$= \frac{1.60 \times 10^{-19}\ \text{coul} \times 4.00 \times 10^{4}\ \text{V} \times 3.11 \times 10^{-11}\ \text{m}}{3.00 \times 10^{8}\ \text{m/seg}}$$

$$= 6.64 \times 10^{-34}\ \text{joule-seg}$$

Esto está en concordancia con el valor de h deducido del efecto fotoeléctrico y del efecto Compton.

Mediciones de V, λ_{min}, y c proporcionan uno de los métodns más precisas para evaluar la relación h/e. En 1951 Bearden, Johnson y Watts en la Universidad Johns Hopkins encontraron usando este procedimiento $h/e = 1.37028 \times 10^{015}$ joule-seg/coul. Esta relación se combina con muchas otras medidas combinadas de constantes físicas, el conjunto de datos se analiza mediante métodos estadísticos elaborados para encontrar el "mejor" valor para las diferentes constantes físicas. Los mejores valores cambian (pero usualmente sólo dentro de una precisión estimada *a priori*) y se hacen más precisos conforme se tienen mejores datos experimentales. ◀

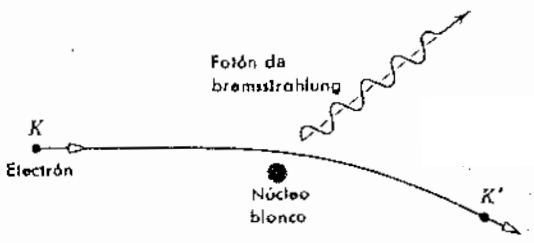


FIGURA 2-11
El proceso de bremsstrahlung responsable de la emisión de rayos X en el espectro continuo.

2.7 Producción y aniquilación de pares

Además de los efectos fotoeléctrico y Compton existe otro proceso mediante el cual los fotones pierden su energía e interacción con la materia, este proceso es el llamado *producción de pares*. La producción de pares también es un ejemplo excelente de la conversión de energía radiante en energía de masa en reposo y energía cinética. En este proceso que se ilustra esquemáticamente en la figura 2-12 un fotón de alta energía pierde toda su energía $h\nu$ en una colisión con un núcleo creando un electrón y un positrón (*el par*) y proporcionándoles energía cinética. Un *positrón* es una partícula que es idéntica en todas sus propiedades a un electrón, excepto en el signo de su carga (y en su momento magnético) que es opuesta a la del electrón; un positrón es un electrón cargado positivamente. En la producción de pares la energía que toma el retroceso del núcleo es despreciable debido a que es muy masivo y por lo tanto el balance de la energía relativista total en el proceso es simplemente

$$h\nu = E_- + E_+ = (m_0c^2 + K_-) + (m_0c^2 + K_+) = K_- + K_+ + 2m_0c^2 \quad (2\text{-}15)$$

En esta expresión E_- y E_+ son las energías relativistas totales y K_- y K_+ son las energías cinéticas del electrón y positrón, respectivamente. Ambas partículas tienen la misma energía de masa en reposo m_0c^2. El positrón se produce con una energía cinética ligeramente mayor a la del electrón debido a que la interacción de Coulomb del par con el núcleo cargado positivamente produce una aceleración en el positrón y una desaceleración en el electrón.

Al analizar este proceso se ignoran los detalles de la interacción misma, considerando solamente la situación antes y después de la interacción. Los principios que nos guían son la conservación de la energía relativista total, la conservación del impulso y la conservación de la carga. De estas leyes de conservación no es difícil demostrar que un fotón no puede simplemente desaparecer en el espacio vacío creando un par al momento que desaparece. La presencia de un núcleo masivo (el cual puede absorber impulso sin afectar apreciablemente el balance de energía) es necesaria para permitir que tanto la energía como el impulso se conserven en el proceso. La carga se conserva automáticamente ya que el fotón no tiene carga y el par de partículas creadas no tienen carga neta. De la ecuación 2-15 se ve que la energía mínima, o umbral, necesaria para que un fotón cree un par es $2m_0c^2$ o 1.02 MeV (1 MeV = 10^6 eV), lo cual es una longitud de onda 0.012 Å. Si la longitud de onda es menor que ésta, correspondiendo a una energía mayor que el valor umbral, el fotón crea el par con energía cinética y con energía en reposo. El fenómeno de producción de pares es un fenómeno de alta energía, los fotones pertenecen a las regiones de rayos X cortos o rayos γ en el espectro electromagnético (ver figura 2-4), región en la cual las energías son iguales o mayores que $2m_0c^2$. Como se verá en la siguiente sección, los resultados experimentales demuestran que la absorción de fotones en la interacción con la materia ocurre principalmente mediante el proceso fotoeléctrico para energías bajas, por el efecto Compton para energías medias y mediante producción de pares para energías altas.

Los pares electrón-positrón se producen en la naturaleza por fotones de los rayos cósmicos y en el laboratorio por fotones del bremsstrahlung provenientes de los aceleradores de partículas.

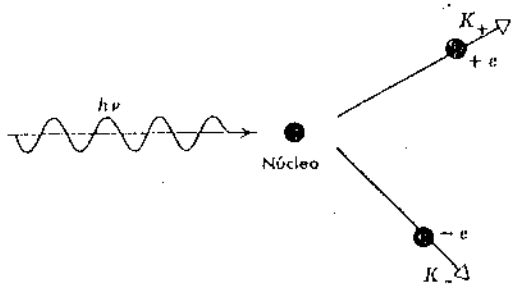


FIGURA 2-12
Proceso de producción de pares.

También se pueden producir otros pares de partículas tales como el protón y antiprotón si el fotón inicial tiene la energía suficiente. Debido a que el electrón y el positrón tienen la masa en reposo más pequeña entre las partículas conocidas, la energía umbral de su producción es la menor. Los experimentos demuestran la descripción cuántica del proceso de producción de pares, mientras que no existe explicación satisfactoria de este fenómeno en la teoría clásica.

Ejemplo 2-6. El análisis de una fotografía de una cámara de burbujas revela la creación de un par electrón-positrón cuando los fotones pasan a través de la materia. Las trazas del electrón y del positrón tienen curvaturas opuestas en el campo magnético uniforme B de 0.20 weber/m^2 y sus radios r son de 2.5×10^{-2} m. ¿Cuál era la energía y longitud de onda del fotón productor del par?

El impulso p del electrón está dada por

$$p = eBr = 1.6 \times 10^{-19} \text{ coul} \times 2.0 \times 10^{-1} \text{ weber/m}^2 \times 2.5 \times 10^{-2} \text{ m}$$
$$= 8.0 \times 10^{-22} \text{ kg-m/seg}$$

Su energía relativista total está dada por

$$E_-^2 = c^2p^2 + (m_0c^2)^2$$

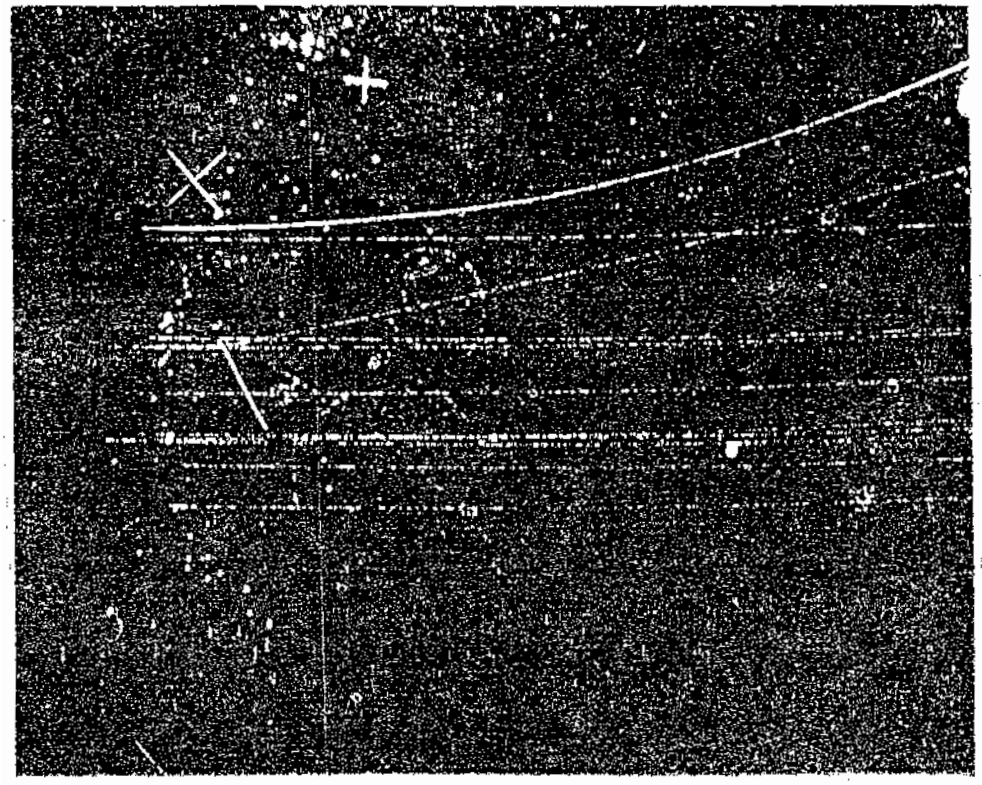


FIGURA 2-13

Producción de un par de electrones visibles en una cámara de burbujas. Las trazas del electrón y del positrón son las dos espirales que se ven en el punto en el que tiene lugar la producción en el líquido que llena la cámara. El estudiante puede determinar cuál de las dos espirales pertenece al positrón sabiendo que las trazas largas son inicialmente deuterones cargados positivamente y que incidió desde la izquierda. (Cortesía de C. R. Sun, Universidad Estatal de Nueva York en Albany).

Puesto que $m_0c^2 = 0.51$ MeV y $pc = 8.0 \times 10^{-22}$ kg-m/seg $\times 3.0 \times 10^8$ m/seg $= 2.4 \times 10^{-13}$ joule $= 1.5$ MeV, se tiene $E^2 = (1.5 \text{ MeV})^2 + (0.51 \text{ MeV})$ y $E = 1.6$ MeV.

La energía relativista total del positrón tenía el mismo valor puesto que su traza tenía el mismo radio, de modo que la energía del fotón era

$$h\nu = E_- + E_+ = 3.2 \text{ MeV}$$

La longitud de onda del fotón se sigue de

$$E = h\nu = hc/\lambda$$

o

$$\lambda = \frac{hc}{E} = \frac{6.6 \times 10^{-34} \text{ joule-sec} \times 3.0 \times 10^8 \text{ m/sec}}{3.2 \times 10^6 \text{ eV} \times 1.6 \times 10^{-19} \text{ joule/eV}} = 3.9 \times 10^{-13} \text{ m} = 0.0039 \text{ Å} \quad \blacktriangleleft$$

Intimamente relacionado con la producción de pares está el proceso inverso llamado aniquilación de pares. Un electrón y un positrón los cuales están inicialmente en reposo y cerca entre sí, se unen y se aniquilan. La materia desaparece y en su lugar se tiene energía radiante. Puesto que el impulso inicial del sistema es cero y el impulso se debe conservar en el proceso, no es posible tener un solo fotón creado ya que un solo fotón no puede tener impulso cero. El proceso más probable es la creación de dos fotones moviéndose con impulsos iguales pero en dirección opuesta. Menos probable, pero posible, es la creación de tres fotones.

En el proceso de dos fotones ilustrado por la figura 2-14, la conservación del impulso da $0 = P_1 + P_2$ ó $P_1 = -P_2$ de modo que los impulsos de los fotones son opuestos en dirección pero iguales en magnitud. Entonces $p_1 = p_2$ o $h\nu_1/c$ y $\nu_1 = \nu_2 = \nu$. Por lo tanto, la conservación de la energía relativista total requiere que $m_0c^2 + m_0c^2 = h\nu + h\nu$, el positrón y el electrón no tienen energía cinética inicial y las energías de los fotones son las mismas. En consecuencia, $h\nu = m_0c^2 = 0.51$ MeV, correspondiendo a un fotón de longitud de onda de 0.024 Å. Si el par inicial tuviera alguna energía cinética, entonces la energía del fotón excedería los 0.51 MeV y su longitud de onda debería ser menor que 0.024 Å.

En el proceso de producción de pares se crean los positrones, los cuales al pasar a través de la materia pierden energía en colisiones sucesivas hasta que se combinan con un electrón para formar un sistema ligado llamado *positronio*. El "átomo" positronio es de vida corta decayendo en fotones a los 10^{-10} segundos de su formación. Supuestamente, el electrón y el positrón se mueven alrededor de su centro de masa común en una especie de danza de la muerte antes de su aniquilación mutua.

Ejemplo 2-7. (a) Supóngase que la figura 2-14 representa el proceso de aniquilación en un sistema de referencia S en el cual el par electrón-positrón está en reposo y los dos fotones de la aniquilación moviéndose a lo largo del eje x. Encontrar la longitud de onda λ de estos fotones en términos de m_0, la masa en reposo de un electrón o positrón.

Se ha visto que p_1 es igual a p_2 y $h\nu_1$ igual a $h\nu_2$. Cada uno de los fotones tiene la misma energía, la misma frecuencia y la misma longitud de onda. Por lo tanto se pueden suprimir los subíndices, y de las relaciones $h\nu = m_0c^2$ y $p = E/c$ se obtiene

$$p = E/c = h\nu/c = m_0c^2/c = m_0c$$

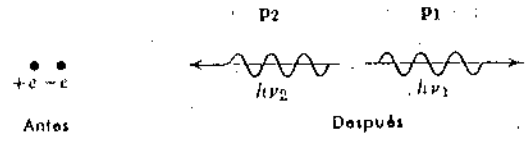


FIGURA 2-14
Aniquilación de un par que produce dos fotones.

Pero también se tiene la relación

$$p = h/\lambda$$

de modo que

$$\lambda = h/p = h/m_0c$$

Por lo tanto, en el sistema en el cual el átomo positronio está en reposo cada uno de los fotones tiene la misma longitud de onda $\lambda = h/m_0c$.

(b) Considérese ahora el mismo evento de aniquilación observado en el sistema S', que se mueve en relación a S con una velocidad v y hacia la izquierda. ¿Qué longitud de onda asignará este observador (móvil) a los fotones de aniquilación?

Aquí, inicialmente el par tiene la energía total relativista $2mc^2$, donde m es la masa relativista, en vez de tener solamente la energía con masa en reposo $2m_0c^2$, así que la conservación de la energía en el proceso de aniquilación conduce a

$$2mc^2 = p_1'c + p_2'c$$

Al mismo tiempo, ahora el par se mueve con velocidad v a lo largo de la dirección positiva del eje x' de modo que su impulso inicial es $2mv$ en vez de ser cero como antes. Ahora la conservación del impulso produce

$$2mv = p_1' - p_2'$$

y los fotones se mueven en direcciones opuestas a lo largo del eje x'. Combinando estas dos expresiones, multiplicando la segunda por c y sumándola a la primera se obtiene,

$$p_1' = m(c + v) = \frac{m_0(c + v)}{\sqrt{1 - v^2/c^2}} = m_0c\sqrt{\frac{c + v}{c - v}}$$

Considerando que $p_1' = h/\lambda_1'$, así que

$$\lambda_1' = \frac{h}{p_1'} = \frac{h}{m_0c}\sqrt{\frac{c - v}{c + v}} = \lambda\sqrt{\frac{c - v}{c + v}} \qquad (2\text{-}16a)$$

De manera similar, restando la segunda ecuación de la primera, se obtiene

$$\lambda_2' = \frac{h}{p_2'} = \frac{h}{m_0c}\sqrt{\frac{c + v}{c - v}} = \lambda\sqrt{\frac{c + v}{c - v}} \qquad (2\text{-}16b)$$

Los fotones *no* tienen la misma longitud de onda, pero *se corren por efecto Doppler* de una longitud de onda λ que tenían en el sistema en reposo de la fuente (el átomo positrónio). Si un observador está situado en el eje x de manera que la fuente se mueve *hacia él*, recibirá el fotón 1 con una frecuencia *mayor* que la frecuencia en "reposo". Si un observador está situado en el eje x' de tal modo que la fuente se mueve *alejándose de él*, recibirá el fotón 2 con una frecuencia *menor* que la frecuencia en reposo. En realidad, este ejemplo es una derivación de la fórmula del *corrimiento Doppler longitudinal* de la teoría de la relatividad. ◀

La primera evidencia experimental del proceso de producción de pares y de la existencia de los positrones fue obtenida por Anderson en 1933 mientras investigaba la radiación cósmica. Esta radiación consiste en un flujo de fotones con gran contenido energético y partículas cargadas que inciden sobre la tierra provenientes de fuentes extraterrestres. Anderson utilizó una cámara de niebla que contenía una placa delgada de plomo, encontrándose todo el aparato en un campo magnético. Al exponer este aparato a la radiación cósmica se encontró que de una manera irregular salían desde algún punto de la placa de plomo un par de partículas cargadas. Se supuso que estos eventos eran el resultado de la interacción de un fotón con el plomo ya que no se había visto que una partícula cargada incidiera en el punto de expulsión,

mientras que un fotón por no poseer carga podría haber incidido en el punto de expulsión sin ser visto. Las dos partículas cargadas expulsadas en estos eventos eran desviadas por el campo magnético en direcciones opuestas, de tal manera que sus cargas eran de signo opuesto. De otras consideraciones se puede mostrar que las magnitudes de estas cargas son igual a una carga electrónica y que las masas de las partículas son aproximadamente igual a una masa electrónica.

El descubrimiento del proceso de la producción de pares vino a explicar el origen de una discrepancia entre la, entonces usual, teoría de la atenuación de los rayos X y las medidas de los coeficientes de atenuación para los rayos X de 2.6 MeV (rayos obtenidos de una fuente radiactiva) por diversos materiales. Como originalmente la teoría no incluía la producción de pares, la atenuación predicha era muy pequeña; con la inclusión del proceso de la producción de pares el acuerdo que se obtiene entre la teoría y el experimento es bueno. Sin embargo, la importancia real del descubrimiento de Anderson radica en la precisa confirmación que éste vino a dar a la teoría relativista mecánico-cuántica del electrón de Dirac.

La teoría de Dirac conduce a la predicción de que los valores permitidos de la energía relativista total E para un electrón libre son

$$E = \pm\sqrt{c^2p^2 + (m_0c^2)^2} \tag{2-17}$$

donde m_0 es la masa en reposo del electrón. Estos valores son simplemente las soluciones para E de la ecuación (2-6), pero la solución con el signo menos corresponde a una energía total relativista negativa, concepto tan advenedizo para la mecánica relativista como lo es la energía total negativa para la mecánica clásica. En vez de desechar la parte negativa sobre la base de su falta de realidad física, Dirac escudriñó las consecuencias de la ecuación completa y al hacerla llegó a algunas conclusiones muy interesantes. Considérese la figura 2-15 que es un diagrama de niveles de energía que representa la ecuación (2-17). Si existe el espectro continuo de niveles de energía que se indica, todos los electrones libres de energía positiva estarían disponibles para realizar transiciones entre estos niveles, transiciones acompañadas por la emisión de fotones con las energías apropiadas: esto obviamente no concuerda con el experimento, debido a que por lo general no se observa que electrones libres emitan espontáneamente fotones de energía $h\nu \geq 2m_0c^2$. Sin embargo, Dirac señaló que se puede superar esta dificultad si se supone que todos los niveles de energía negativa están llenos normalmente en todos los puntos del espacio. De acuerdo con esta suposición, *un vacío consiste de un mar de electrones en niveles de energía negativa*. Esto sí concuerda con el experimento. Por ejemplo, la carga negativa no podría ser detectada si se la supone uniformemente distribuida y por lo tanto no ejercería fuerza sobre un cuerpo cargado. Consideraciones similares demostrarán que todas las propiedades "usuales" de un mar de electrones de energía negativa son tales que su presencia no sería aparente en ninguna de los experimentos usuales. Empero, la teoría de Dirac del vacío

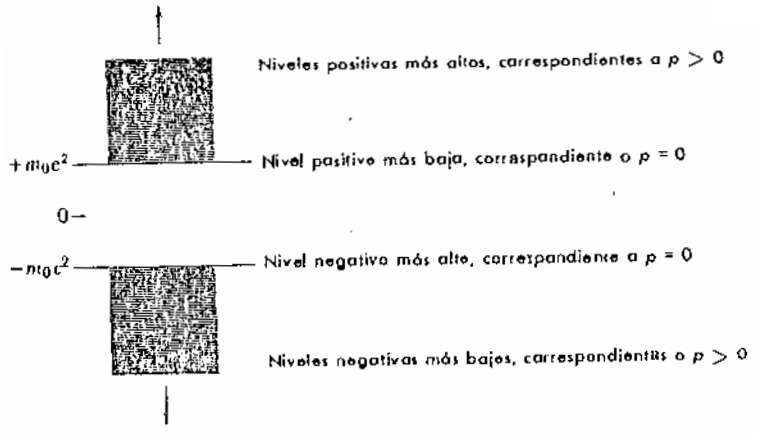


FIGURA 2-15
Niveles de energía de un electrón libre según la teoría de Dirac.

no es completamente vacía ya que predice ciertas propiedades nuevas que pueden comprobarse por el experimento.

El diagrama de niveles de energía para un electrón libre sugiere la posibilidad de excitar un electrón que se encuentra en un nivel de energía negativa mediante la absorción de un fotón. Puesto que todos los niveles de energía negativa se consideran totalmente ocupados, el electrón deberá ser excitado a uno de los niveles de energía positiva desocupados. Obviamente, la energía mínima necesaria para este proceso es $h\nu = 2m_0c^2$ y el resultado del proceso es la producción de un electrón en un nivel de energía positiva más un agujero en un nivel de energía negativa. Se puede demostrar que un agujero en un nivel de energía negativa tiene todas las propiedades mecánicas y ópticas de un positrón de energía positiva. Por ejemplo, existe una carga positiva $+e$ asociada con la ausencia de un electrón con carga negativa $-e$. Consecuentemente, este es el proceso de producción de pares observado experimentalmente por Anderson tres años después de que fuera predicho teóricamente por Dirac.

2.8 Secciones transversales para absorción y dispersión de fotones

Considérese, como en la figura 2-16 un haz paralelo de fotones que se pasan a través de una placa material. Los fotones pueden interactuar con los átomos de la placa mediante cuatro procesos diferentes: fotoeléctrico, producción de pares, Thomson y Compton. En los dos primeros los fotones se absorben completamente, mientras que en los dos últimos sólo son dispersados, pero en todos los procesos ocurre la eliminación de fotones del haz. La pregunta de qué tan probable es que estos procesos ocurran, para un conjunto de circunstancias dadas, es una de las preguntas de gran importancia tanto teóricas como prácticas. Por ejemplo, para un físico médico es muy importante diseñar el blindaje para una máquina de rayos X, o para un ingeniero nuclear diseñar el blindaje para un reactor. La respuesta a esta pregunta se expresa en términos de las cantidades llamadas secciones transversales. En este libro las secciones transversales aparecen primero en conexión con los fotones pero posteriormente aparecerán en conexión con otras cosas.

La probabilidad de que un fotón de una energía dada sea, por ejemplo, absorbido en un proceso fotoeléctrico al pasar por un átomo de la placa queda especificada por el valor de la *sección transversal* fotoeléctrica $\sigma_{f\cdot E}$. Esta medida de la probabilidad de ocurrencia del proceso fotoeléctrico se define de modo que el número $N_{f\cdot E}$ de absorciones fotoeléctricas que ocurren es

$$N_{f\cdot E} = \sigma_{f\cdot E} I n \tag{2-18}$$

cuando un haz que contiene I fotones incide sobre una placa cortada que contiene n átomos por unidad de área. Aquí se supone que la placa es lo suficientemente delgada como para que la

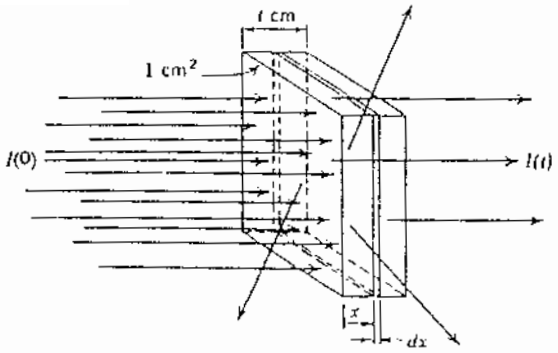


FIGURA 2-16

Haz de fotones atravesando una lámina.

probabilidad de que un fotón dado sea absorbido al pasar por la placa sea mucho menor que la unidad.

La definición de (2-18), que es un prototipo de las definiciones de todas las secciones transversales, es suficientemente importante para garantizar una interpretación física prudente. Obsérvese primero que el número N_{PE} de absorción en efecto debería incrementarse en proporción al número I de fotones que inciden sobre la placa. Además, si la placa es delgada en el sentido en el que se especificó anteriormente, los átomos en la placa no se "ensombrecerán" entre sí, por lo que toca a los fotones incidentes. Entonces, el número N_{PE} de absorciones también se incrementará en proporción al número n de átomos blanco por unidad de área de la placa; por lo que se tendrá

$$N_{PE} \propto In$$

Si esta proporcionalidad se escribe como una igualdad en la que a la constante de proporcionalidad se la denota por σ_{PE}, se obtiene la ecuación que define la sección transversal. Por lo tanto se ve que la sección transversal cuyo valor depende tanto de la energía del fotón como del tipo de átomo, es una medida de cuán efectivos son esos átomos para absorber los fotones mediante el efecto fotoeléctrico. Puesto que las cantidades N_{PE} e I en la ecuación 2-18 son adimensionales mientras que n tiene dimensiones de (área)$^{-1}$, es claro que σ_{PE} deberá tener dimensiones de (área). Por lo tanto, resulta razonable usar el nombre de sección transversal para σ_{PE}. Casi es posible dar una interpretación geométrica si se imagina un círculo de área σ_{PE} centrado en cada átomo en la placa con la propiedad de que cualquier fotón que entre a esta placa circular es absorbido por el átomo mediante el efecto fotoeléctrico. Esta interpretación geométrica resulta conveniente para la visualización y aún para los cálculos pero definitivamente no deberá considerarse literalmente cierta. En realidad una sección transversal es simplemente una forma de expresar numéricamente la probabilidad de que cierto tipo de átomos origine un fotón de una energía dada para que realice un proceso particular. Las definiciones e interpretaciones de las secciones transversales para los otros procesos de absorción o dispersión son completamente análogas a las que se han considerado en este ejemplo.

La figura 2-17 muestra las secciones transversales de dispersión (σ_S), producción de pares (σ_{PR}), fotoeléctrica (σ_{PE}), y total (σ) medidas para un átomo de plomo como una función de la energía $h\nu$ del fotón. La sección transversal de dispersión especifica la probabilidad de que la dispersión ocurra por cualquiera de los procesos Thomson y Compton. Para el plomo que tiene un número atómico grande y por lo tanto electrones atómicos fuertemente ligados la dispersión Thomson domina sobre la dispersión Compton cuando la energía del fotón está por debajo de $h\nu = 10^5$ eV. El rompimiento brusco en la sección transversal fotoeléctrica ocurre para las energías de enlace de los diferentes electrones en el átomo de plomo; cuando $h\nu$ cae por debajo de la energía de enlace de un electrón particular ya no es posible energéticamente un proceso fotoeléctrico en el que él intervenga. La sección transversal para producción de pares crece muy rápidamente desde cero cuando $h\nu$ excede la energía umbral $2m_0c^2 \simeq 10^6$ eV necesaria para materializar el par. En la figura 2-17 la sección transversal total σ es la suma de las secciones transversales de dispersión, fotoeléctrica y de producción de pares; esta cantidad especifica la probabilidad de que un fotón tenga cualquier tipo de interacción con el átomo. De la figura se pueden ver los intervalos de valores de la energía en los cuales cada uno de los tres procesos hace su contribución más importante, para el caso del plomo:

Efecto fotoeléctrico:	$h\nu < 5 \times 10^5$ eV
Dispersión:	5×10^5 eV $< h\nu < 5 \times 10^6$ eV
Producción de pares:	5×10^6 eV $< h\nu$

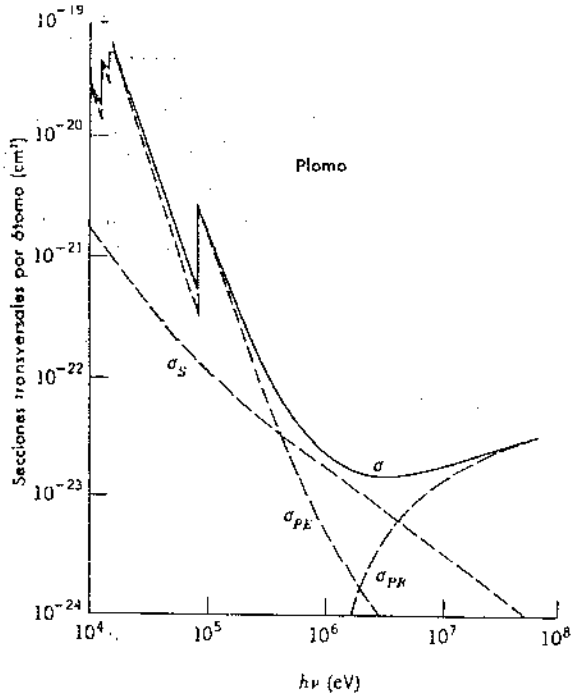


FIGURA 2-17

Secciones transversales de dispersión, fotoeléctrica, de producción de pares y total para un átomo de plomo.

Debido a que estos procesos tienen probabilidades diferentes dependiendo del número atómico, los intervalos de energía en los cuales ellos dominan son muy distintos para átomos con número atómico bajo. Para el aluminio, los intervalos de energía son aproximadamente:

Efecto fotoeléctrico: $h\nu < 5 \times 10^4$ eV
Dispersión: 5×10^4 eV $< h\nu < 1 \times 10^7$ eV
Producción de pares: 1×10^7 eV $< h\nu$

Ejemplo 2-8. Evaluar la atenuación de un haz de rayos X paralelos al pasar a través de una placa *gruesa* de materia en términos de la sección transversal total.

Con referencia a la figura 2-16, cuando el haz incide sobre la cara frontal de la placa de espesor t el número de fotones en el haz es $I(0)$ y la sección cortada contiene ρ átomos por cm³. Supóngase, por simplicidad, que el área de la sección cortada es de 1 cm². Debido a los procesos de dispersión y absorción, el haz paralelo contiene un número pequeño $I(x)$ de fotones después de haber penetrado en la sección cortada x cm. Considérese una lámina delgada de la placa de espesor dx y localizada en x. El número de átomos por cm² en la lámina es ρ veces su volumen dx o ρdx. En una definición análoga a (2-18), el número de fotones del haz que serán dispersados o absorbidos en la lámina está especificado por la sección transversal total σ y es $\sigma I(x)\rho\, dx$. De esta manera, el número de fotones del haz que salen de la lámina, $I(x + dx)$, que es igual al número incidente menos el número que se ha perdido, es

$$I(x + dx) = I(x) - \sigma I(x)\rho\, dx$$

$$dI(x) \equiv I(x + dx) - I(x) = -\sigma I(x)\rho\, dx$$

Para encontrar el número $I(t)$ de fotones que salen por la cara posterior de la placa se despeja $dI(x)/I(x)$ y se integra sobre x

$$\begin{aligned}
\frac{dI(x)}{I(x)} &= -\sigma\rho\, dx \\
\int_0^t \frac{dI(x)}{I(x)} &= -\sigma\rho \int_0^t dx \\
\ln I(x)\Big]_0^t &= -\sigma\rho\, t \\
\ln \frac{I(t)}{I(0)} &= -\sigma\rho\, t \\
\frac{I(t)}{I(0)} &= e^{-\sigma\rho t} \\
I(t) &= I(0)e^{-\sigma\rho t}
\end{aligned} \tag{2-19}$$

La intensidad del haz, medida por el número I de fotones que contiene, decrece exponencialmente conforme el espesor t de la placa aumenta. La cantidad $\sigma\rho$, llamada *coeficiente de atenuación*, tiene dimensiones de (cm^{-1}) y es el recíproco del espesor de la placa necesaria para atenuar la intensidad del haz por un factor e. A este espesor se le llama la *longitud de atenuación* Λ y es

$$\Lambda = 1/\sigma\rho \tag{2-20}$$

Por supuesto, el coeficiente de atenuación tiene la misma dependencia con la energía del fotón que la sección transversal total. La figura 2-18 muestra medidas de los coeficientes de atenuación para el plomo, el estaño y el aluminio para fotones de energía relativamente alta.

Esta sección resume muchos de los aspectos prácticos de los fenómenos de emisión y absorción de radiación electromagnética que se han estudiado en este capítulo. Pero los aspectos fundamentales de estos fenómenos se resumen mejor afirmando que demuestran que la radiación electromagnética está cuantizada en partículas energéticas llamadas fotones. También deberá darse énfasis al hecho de que los fenómenos de interferencia y difracción demuestran que los fotones no viajan desde un sistema en el que son emitidos hasta donde son absorbidos en la forma simple en la que lo hacen las partículas clásicas: sino que los fotones actúan como si fueran guiados por ondas clásicas debido a que al viajar a través de un sistema, tal como un aparato de difracción, quedan mejor descritos por ondas clásicas que se propagan en dicho aparato.

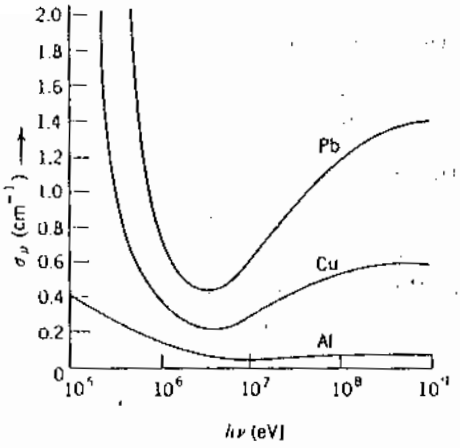


FIGURA 2-18

Coeficientes de atenuación para distintos átomos para un intervalo de energía del fotón.

PREGUNTAS

1. En los experimentos de efecta fotocléctrico, la corriente (número de electrones emitidos por unidad de tiempo) es proporcianal a la intensidad de la luz. ¿Se puede usar solamente este resultada para distinguir entre las teorías clásica y cuántica?

2. ¿Por qué en la figura 2-2 la corriente fatoeléctrica no aumenta verticalmente hacia su máximo valor (saturación) cuando la diferencia de potencial aplicada es ligeramente más pasitiva que $-V_s$?

3. ¿Por qué, aun cuando la radiación incidente es monocromática los fatoelectranes son emitidos con una gama de velocidades?

4. A la existencia de una frecuencia de carte en el efecta fotocléctrico se le considera can frecuencia como la nbjeción más fuerte a la teoría ondulataria. Explique.

5. ¿Por qué las mediciones en el efecto fotoeléctrico son muy sensibles a la superficie fotocléctrica?

6. ¿Los resultados de los experimentos fotaeléctricos invalidan el experimento de interferencia de Young?

7. ¿Se puede utilizar el artificio de hacer $h \rightarrow 0$ para abtener los resultados clásicos a partir de los cuánticos en el caso del efecto fotoeléctrico? Explique.

8. Suponga que la emisión de fotones desde una fuente de radiación es caótica en dirección. ¿Podría esperar que la intensidad (o la densidad de energía) variara inversamente al cuadrado de la distancia de la fuente en la teoría de las fotones como acurre en la teoría ondulatoria?

9. ¿Tiene masa un fotón de energía E? Si es así, evalúela.

10. ¿Por qué en la dispersión de Compton pndría Ud. esperar que $\Delta\lambda$ fuera independiente de los materiales que componen al dispersor?

11. ¿Cuándo esperaría Ud. observar más fácilmente el efecto Compton con blancas dispersores compuestos por átomos con número atómico grande o con átomos con número atómico bajo? Explique.

12. ¿Se observa el efectn Compton con luz visible? ¿Par qué?

13. De la teoría electramagnética clásica, ¿esperaría un mínimo definido en la langitud de onda de la radiación emitida cuando el valor de la energía de un electrón incidente snbre un blanco de un tubo de rayos X está dado?

14. Un tuba de televisián, ¿emite rayos X? Explique.

15. ¿Qué efecto(s) produce, sobre el espectro de rayos X resultante, el decremento del voltaje del tubo?

16. Analizar el proceso de bremsstrahlung como el inversa del proceso Compton y como el inverso del praceso fotoeléctrico.

17. Describir varias métodos que se pueden utilizar para determinar experimentalmente el valor de la constante de Planck (h).

18. ¿De qué factores esperaría Ud. juzgar si un fotón pierde su energía en la interacción con la materia por proceso fotoeléctrico, por procesn Compton o por prneeso de producción de pares?

19. ¿Puede Ud. dar evidencia experimental que contradiga la idea de que el vacío es un mar de electrones en estados de energía negativa?

20. ¿Puede ocurrir una aniquilación electrán-positrón con la creación de *un* fotán, si existe un núcleo cercano disponible para compensar el impulso?

3

Postulado de de Broglie Propiedades ondulatorias de las partículas

3.1 Ondas de materia

Maurice de Broglie era un físico francés quien, desde un principio, apoyó el punto de vista de Compton sobre la naturaleza corpuscular de la radiación. Sus experimentos y estudios sobre los problemas filosóficos de la física de ese tiempo, impresionaron tanto a su hermano Louis, que éste cambió su carrera de historia a física. En 1924, Louis de Broglie propuso la existencia de ondas de materia en la tesis doctoral que presentó a la Facultad de Ciencias de la Universidad de París. La original y concienzudo de su tesis fue reconocido inmediatamente, sin embargo, sus ideas fueron consideradas como carentes de realidad física por su aparente falta de evidencias experimentales. Fue Albert Einstein quien reconoció su importancia y validez y atrajo hacia ella la atención de otros físicos. Cinco años después, de Broglie obtuvo el premio Nobel en física, cuando sus ideas fueron experimentalmente confirmadas en forma dramática.

La hipótesis de de Broglie consistía en que el comportamiento dual de la radiación, es decir, onda-partícula, debería ser igualmente aplicable a la materia. Así como el fotón tiene asociado una onda de luz que gobierna su movimiento, una partícula de materia (por ejemplo, un electrón) tiene asociada una onda de materia que gobierna su movimiento. Ahora bien, como el universo consiste en general, de radiación y materia, la sugerencia de de Broglie es esencialmente una afirmación acerca de una gran simetría de la naturaleza. En efecto, propuso que los aspectos ondulatorios de la materia están relacionados con los aspectos corpusculares en la misma forma cuantitativa que en el caso de la radiación. De acuerdo con de Broglie, tanto para la materia como para la radiación, la energía total E de un ente se relaciona con la frecuencia ν de la onda asociada a su movimiento por medio de la ecuación

$$E = h\nu \tag{3-1a}$$

y el impulso p del ente se relaciona con la longitud de onda λ de la onda asociada por la ecuación

$$p = h/\lambda \tag{3-1b}$$

En lo anterior, los conceptos corpusculares energía E, e impulso P, se relacionan con los conceptos ondulatorios frecuencia ν y longitud de onda λ a través de la constante de Planck h. A la ecuación (3-1b) escrita en la forma siguiente se le denomina *relación de de Broglie*.

$$\lambda = h/p \tag{3-2}$$

que predice la *longitud de onda de de Broglie* de una *onda de materia* asociada con el movimiento de una partícula material de impulso p.

Ejemplo 3-1. (a) ¿Cuál será la longitud de onda de de Broglie de una pelota de béis-bol que se mueve a una velocidad $v = 10$m/seg?

Suponga $m = 1.0$ kg. De (3-2)

$$\lambda = \frac{h}{p} = \frac{h}{mv} = \frac{6.6 \times 10^{-34} \text{ joule-seg}}{1.0 \text{ kg} \times 10 \text{ m/seg}} = 6.6 \times 10^{-35} \text{ m} = 6.6 \times 10^{-25} \text{ Å}$$

(b) ¿Cuál será la longitud de onda de un electrón cuya energía cinética es 100 eV?

En este caso;

$$\lambda = \frac{h}{p} = \frac{h}{\sqrt{2mK}} = \frac{6.6 \times 10^{-34} \text{ joule-seg}}{(2 \times 9.1 \times 10^{-31} \text{ kg} \times 100 \text{ eV} \times 1.6 \times 10^{-19} \text{ joule/eV})^{1/2}}$$

$$= \frac{6.6 \times 10^{-24} \text{ joule-seg}}{5.4 \times 10^{-21} \text{ kg-m/seg}} = 1.2 \times 10^{-10} \text{ m} = 1.2 \text{ Å}$$ ◀

En los experimentos de óptica geométrica, no se manifiesta la naturaleza ondulatoria de la propagación de la luz, ya que, las dimensiones importantes de los aparatos utilizados son muy grandes comparadas con la longitud de onda de la luz. Si se representa por a una dimensión característica de un aparato óptico (es decir, la amplitud de una lente; espejo o rendija) y λ es la longitud de onda de la luz que pasa por el aparato, se está en el dominio de la óptica geométrica siempre que $\lambda/a \to 0$. La razón es que los efectos de difracción en cualquier aparato siempre ocurren para ángulos de aproximadamente $\theta = \lambda/a$, de modo que los efectos de difracción son completamente despreciables cuando $\lambda/a \to 0$. Obsérvese que la óptica geométrica implica la propagación de rayos, que son similares a las trayectorias del movimiento de partículas clásicas.

Sin embargo, cuando la dimensión característica de un aparato óptico a se vuelve comparable con, o menor que, la longitud de onda λ de la luz que pasa por él, se está en el dominio de la óptica física. En este caso $\lambda/a \gtrsim 1$, y el ángulo de difracción $\theta = \lambda/a$ es suficientemente grande como para que los efectos de difracción sean fácilmente observables y la naturaleza ondulatoria de luz resulta aparente. Por lo tanto, para observar características ondulatorias, en el movimiento de materia, se requieren sistemas con aperturas u obstáculos adecuadamente pequeños. Los sistemas de aperturas, en la escala más fina, de que disponían los experimentadores contemporáneos de de Broglie, los constituirían los espacios entre planos de átomos adyacentes en un sólido, donde $a \simeq 1$ Å. (En la actualidad se dispone de sistemas que involucran dimensiones nucleares de $\simeq 10^{-4}$ Å). Considerando los resultados del ejemplo 3-1, se ve que no se puede esperar poder detectar *ningún* comportamiento ondulatorio en el movimiento de una pelota de béis-bol, donde $\lambda/a \simeq 10^{-25}$ para $a \simeq 1$ Å. Sin embargo, para una partícula material con una masa mucho menor que la de una pelota de béis-bol, el impulsn p se reduce y la longitud de onda de de Broglie $\lambda = h/p$ crece lo suficiente como para que puedan observarse efectos de difracción. Utilizando aparatos con dimensiones características $a = 1$ Å, los aspectos ondulatorios en el movimiento del electrón con $\lambda = 1.2$ Å del ejemplo 3-1, resultarían evidentes.

En 1926, Elsasser propuso que la naturaleza ondulatoria de la materia podía ser probada del mismo modo como inicialmente se probó la naturaleza ondulatoria de los rayos X, a saber, haciendo incidir un haz de electrones, con la energía apropiada, sobre un sólido cristalino. Los átomos del cristal son utilizados como un arreglo tridimensional de centros dispersores para la onda electrónica y, por lo tanto, deberán dispersar fuertemente a los electrones en ciertas direcciones características, de igual forma que en la difracción de rayos X. Esta idea fue confirmada experimentalmente por Davisson y Germer en Estados Unidos y por Thomson en Escocia.

El aparato de Davisson y Germer se muestra esquemáticamente en la figura 3-1. Los electrones producidos por un filamento caliente son acelerados a través de una diferencia de potencial V y emergen del "cañón electrónico" C con una energía cinética eV. Este haz de

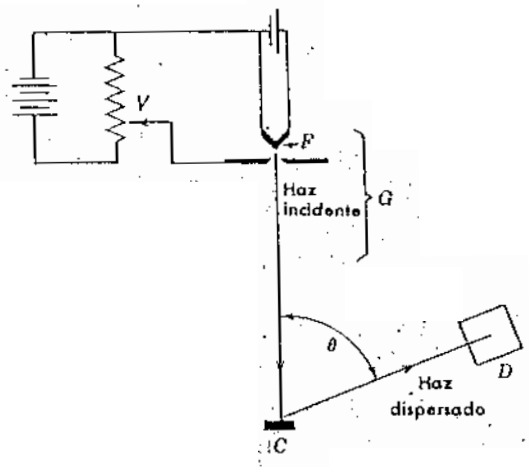


FIGURA 3-1

Aparato de Davisson y Germer. Los electrones de un filamento F son acelerados por una diferencia de potencial variable V. Después de dispersarse en el cristal C, son colectados por el detector D.

electrones incide perpendicularmente sobre un monocristal de nickel en C. El detector D se coloca a un ángulo θ determinado y se realizan medidas de la intensidad del haz dispersado para varios valores del potencial de aceleración V. Por ejemplo, en la figura 3-2, se muestra que un haz fuertemente dispersado es detectado a $\theta = 50°$ para $V = 54V$. La existencia de este "pico" en el patrón de electrones dispersados demuestra cualitativamente la validez del postulado de de Broglie, ya que, *sólo puede explicarse como una interferencia constructiva de ondas dispersadas por el arreglo periódico de los átomos hacia los planos del cristal.* El fenómeno es análogo a las bien conocidas "reflexiones de Bragg" que ocurren en la dispersión de rayos X por los planos atómicos de un cristal y no puede entenderse en base al movimiento de una partícula clásica sino sólo en base al movimiento de ondas. Las partículas clásicas no pueden mostrar interferencia, mientras que las ondas si lo hacen. La interferencia que aquí se implica, no ocurre entre las ondas asociadas con un electrón y las ondas asociadas con otro, sino que se refiere a la interferencia que ocurre entre diferentes partes de la onda asociada con un electrón y que han sido dispersadas por varias regiones del cristal. Esto se puede demostrar, si se utiliza un haz de electrones con una intensidad tan baja que los electrones pasan por un aparato uno por uno y mostrando que el patrón de electrones dispersados permanece invariante.

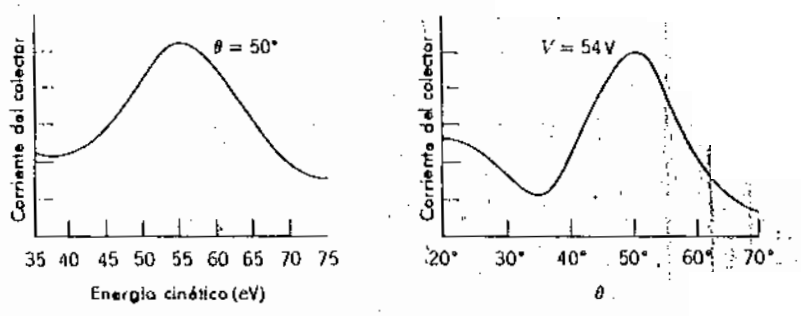


FIGURA 3-2

Izquierda: la corriente del colector en el detector D de la figura 3-1, como función de la energía cinética de los electrones incidentes, mostrando un máximo de difracción. El ángulo θ en la figura 3-1, se ajusta a 50°. Si se utiliza un valor apreciablemente mayor o menor, el máximo de difracción desaparece. *Derecha:* la corriente como función del ángulo del detector para un valor fijo de la energía cinética electrónica, 54 eV.

El origen de una reflexión de Bragg se muestra en la figura 3-3 y la *relación de Bragg*, a que está sujeta, se deriva en el pie de figura.

$$n\lambda = 2d \operatorname{sen}\varphi \tag{3-3}$$

Para las condiciones mostradas en la figura 3-3, la distancia interplanar efectiva d se puede medir por dispersión de rayos X del mismo cristal y resulta ser 0.91 Å. Como $\theta = 50°$ entonces $\varphi = 90° - 50°/2 = 65°$. La longitud de onda calculada de (3-3), suponiendo $n = 1$, resulta

$$\lambda = 2d \operatorname{sen}\varphi = 2 \times 0.91 \text{ Å} \times \operatorname{sen} 65° = 1.65 \text{ Å}$$

Para electrones de 54 eV, la longitud de onda de de Broglie calculada de (3-2), es

$$\lambda = h/p = 6.6 \times 10^{-34} \text{ joule-seg}/4.0 \times 10^{-24} \text{ kg-m/seg} = 1.65 \text{ Å}$$

Esta impresionante concordancia proporciona una confirmación cuantitativa de la relación de de Broglie entre λ, p y h.

La amplitud del pico observado en la figura 3-2, también se puede entender fácilmente, ya que, los electrones de baja energía no pueden penetrar profundamente en el cristal, de modo que sólo un número pequeño de planos atómicos contribuyen a la onda difractada. Por lo tanto, el máximo correspondiente a la difracción no es *estrecho*. Sin lugar a dudas, todos los resultados experimentales estuvieron en excelente concordancia tanto cualitativa como cuantitativa-

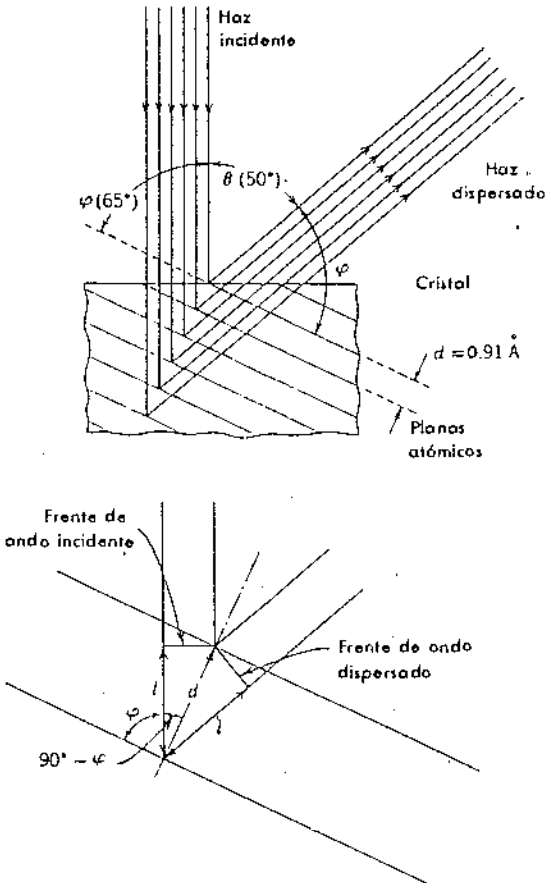


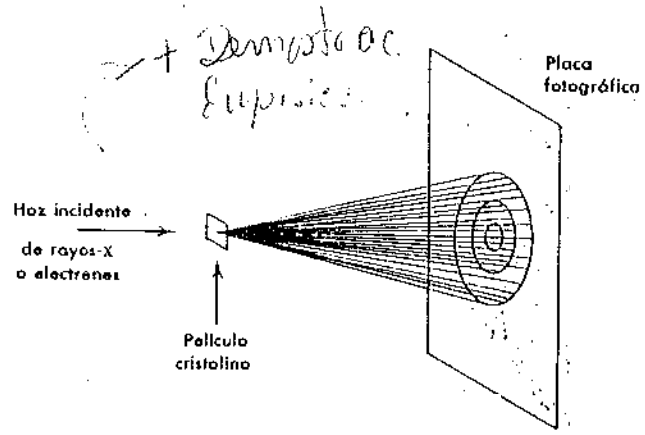
FIGURA 3-3

Arriba: el haz fuertemente difractado a $\theta = 50°$ y $V = 54$V resulta de la dispersión de tipo ondulatorio por la familia de planos atómicos que se muestran, los cuales tienen una distancia de separación $d = 0.91$ Å. El ángulo de Bragg es $\varphi = 65°$. La refracción del haz dispersado al atravezar la superficie del cristal, no se indica por simplicidad. *Abajo:* derivación de la relación de Bragg, mostrando sólo dos planos atómicos y dos rayos correspondientes a los haces incidente y difractado. Si el número de longitudes de onda que caben en la distancia $2l$ entre los frentes de onda incidente y dispersado, es un entero dado por $n\lambda$, medida sobre el rayo inferior, entonces, las contribuciones de ambos rayos al frente de onda dispersada, estarán en fase y se obtendrá un máximo de difracción en el ángulo φ. Como $l/d = \cos(90° - \varphi) = \operatorname{sen}\varphi$, se tiene $2l = 2d \operatorname{sen}\varphi$ y por lo tanto se obtiene la relación de Bragg $n\lambda = 2d \operatorname{sen}\varphi$. El máximo de difracción de "primer orden", ($n = 1$), por lo general el más intenso.

mente con la predicción de de Broglie, aportando evidencias convincentes de que las partículas materiales se mueven de acuerdo con las leyes del movimiento de ondas.

En 1927, G.P. Thomson demostró la difracción de haces electrónicos que pasan a través de películas delgadas, confirmando independientemente y en detalle, la relación de de Broglie $\lambda = h/p$. Así como el experimento de Davisson-Germer es análogo al de Laue en difracción de rayos X (reflexión del arreglo regular de planos atómicos en un monocristal grande), el experimento de Thomson es análogo al método de Debye-Hull-Scherrer de difracción de rayos X por un polvo (transmisión a través de un agregado de cristales pequeños orientados aleatoriamente. Thomson utilizó electrones de mayor energía, que son mucho más penetrantes, de modo que son varios cientos de planos atómicos los que contribuyen a la onda difractada. El patrón de difracción resultante tiene una estructura muy precisa. A manera de comparación, en la figura 3-4 se muestra un patrón de difracción de rayos X y un patrón de difracción de electrones de substancia policristalina (substancia en la que se tiene un gran número de cristales microscópicos orientados aleatoriamente).

Resulta interesante hacer notar que J.J. Thomson en 1897 descubrió el electrón (caracterizándola como una partícula con una relación carga-masa bien definida) y quien recibió el premio Nobel en 1906; fue el



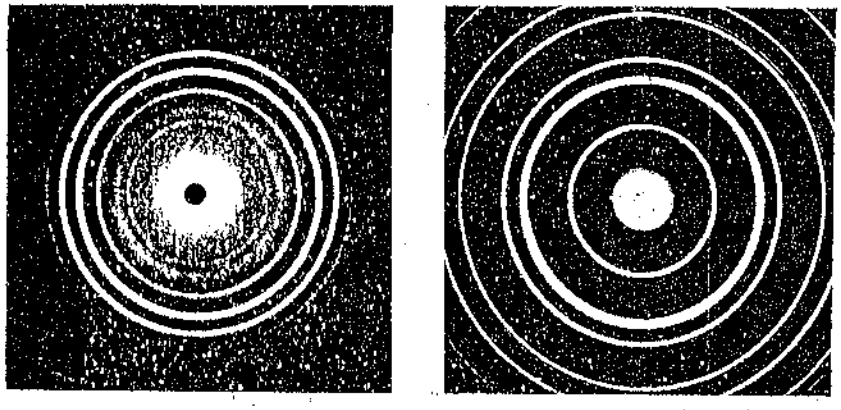


FIGURA 3-4

Arriba: arreglo experimental para la difracción de Debye-Scherrer de rayns X o electrones por un material policristalina. *Abajo a la izquierda:* patrón Debye-Scherrer de difracción de rayos X por cristales de óxido de Zirconio. *Abajo a la derecha:* patrón Debye-Scherrer de difracción de electrones por cristales de oro.

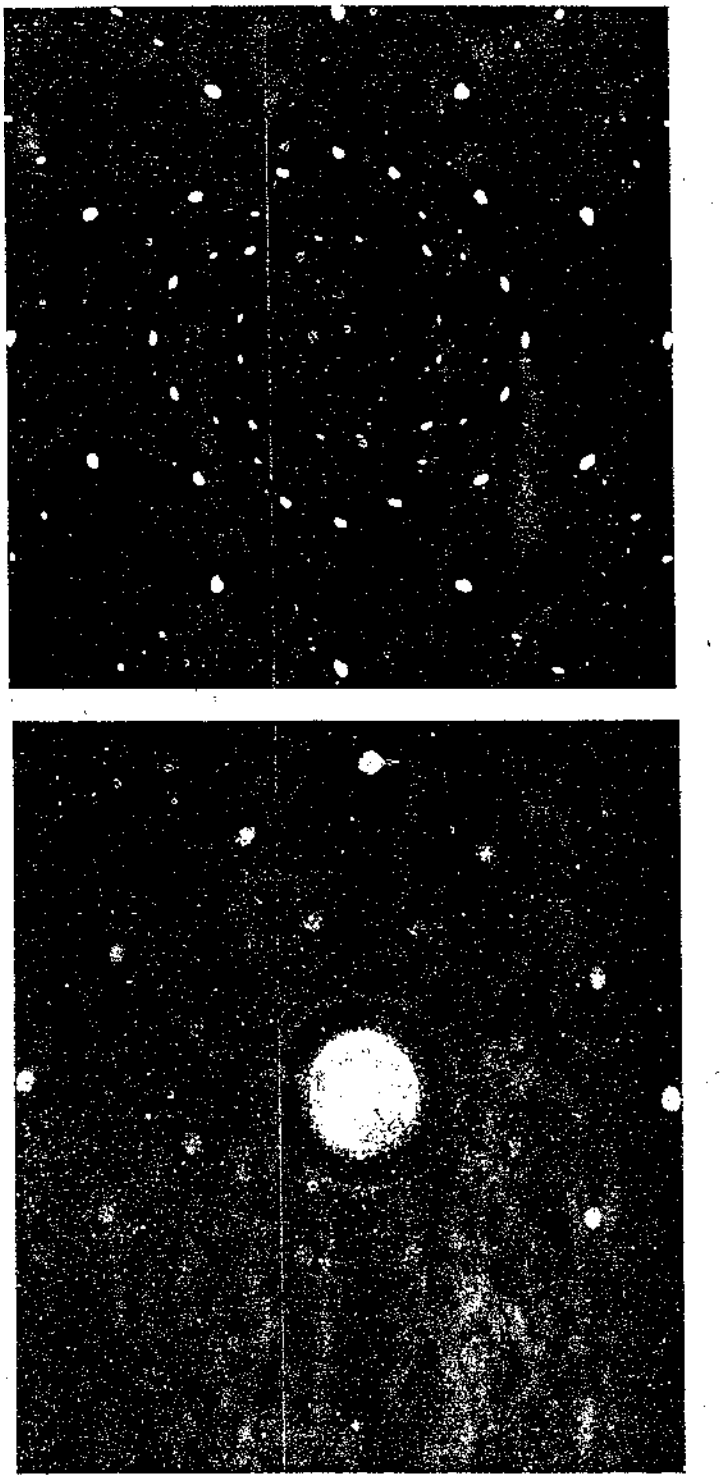


FIGURA 3-5

Arriba: patrón de Laue de la difracción de rayos X por un monocristal de cloruro de sodio. *Abajo:* patrón de Laue de difracción de neutrones de un reactor nuclear por un monocristal de cloruro de sodio.

padre de G.P. Thomson, quien en 1927 descubrió experimentalmente la difracción de electrones y recibió el premio Nobel (con Davisson) en 1937. Acerca de lo anterior escribió Max Jammer: "Uno se inclinaría par decir que a Thomson, el padre, le fue otorgado el premio Nobel por haber demostrado que el electrón es una partícula y a Thomson, el hijo, por haber demostrado que el electrón es una onda".

Todos los objetos materiales, cargados o descargados, no solo los electrones, muestran características ondulatorias en su movimiento, bajo las condiciones de la óptica física. Por ejemplo, Estermann, Stern y Frisch, realizaron experimentos cuantitativos acerca de la difracción de haces moleculares de hidrógeno y haces atómicos de helio, por un cristal de fluoruro de litio; asimismo, Fermi, Marshall y Zinn demostraron la difracción e interferencia de neutrones lentos. En la figura 3-5 se muestra un patrón de difracción de neutrones en un cristal de cloruro de sodio. Inclusive se ha construido un interferómetro que opera con haces de electrones. La existencia de ondas de materia está bien establecida.

Resulta instructivo hacer notar que para poder encontrar evidencias experimentales de la naturaleza ondulatoria de la materia, se tuvo que recurrir a longitudes de onda de de Broglie relativamente grandes. La materia y la radiación, ambas presentan, propiedades de partícula y onda, tanto para longitudes de onda grandes como pequeñas. Los aspectos corpusculares se acentúan, cuando se estudia su emisión o absorción, mientras que los aspectos ondulatorios se acentúan, cuando se estudia su comportamiento al moverse a través de un sistema. Sin embargo, los aspectos ondulatorios de su movimiento son más difíciles de observar a medida que su longitud de onda es más pequeña. Una vez más se puede apreciar el papel tan importante que juega la constante de Planck h. Si h fuera cero, en $\lambda = h/p$ se obtendría $\lambda = 0$ en todos los casos. Por lo tanto, todas las partículas materiales tendrían una longitud de onda menor que cualquier dimensión característica y nunca sería posible observar efectos de difracción. Aunque el valor de h definitivamente no es cero, es pequeño. Esta pequeñez de h es lo que obscurece la existencia de ondas materiales en el mundo macroscópico, ya que, se deberán tener impulsos muy pequeños para poder obtener longitudes de onda mensurables. Para partículas macroscópicas ordinarias, su masa es tan grande que el impulso es siempre lo suficientemente grande como para que la longitud de onda de de Broglie sea suficientemente pequeña y se encuentre fuera del alcance de la detección experimental, y así la mecánica clásica reina suprema. En el mundo microscópico, las partículas materiales tienen masas tan pequeñas que sus impulsos son pequeños aun cuando sus velocidades son bastante altas. Por lo tanto, sus longitudes de onda de de Broglie son lo suficientemente grandes como para ser comparables con las dimensiones características de los sistemas de interés, tales como átomos, y sus propiedades ondulatorias son experimentalmente observables en su movimiento. Sin embargo, no debe olvidarse que en sus interacciones, por ejemplo cuando son detectadas, sus propiedades corpusculares son dominantes aún cuando sus longitudes de onda sean grandes.

Ejemplo 3-2. En los experimentos con átomos de helio a los que anteriormente se hizo referencia, se produce un haz de átomos con velocidad casi uniforme de 1.635×10^5 cm/seg, permitiendo el escape de helio en forma de gas, a través de un pequeño agujero en el recipiente que lo contiene, pasando a una cámara evacuada y luego a través de rendijas muy angostas sobre discos giratorios paralelos de corta separación (un selector mecánico de velocidades). Los átomos inciden sobre la superficie de un cristal de fluoruro de litio y se observa un haz de átomos de helio fuertemente difractado que emerge del cristal. El haz difractado se detectó con un medidor de presión de alta sensibilidad. El análisis usual de los resultados experimentales de la difracción en el cristal indican una longitud de onda de 0.600×10^{-8} cm. ¿Cómo es la concordancia de éste con el valor calculado de la longitud de onda de de Broglie?

La masa de un átomo de helio es:

$$m = \frac{M}{N_a} = \frac{4.00 \text{ g/mol}}{6.02 \times 10^{23} \text{ átom/mol}} = 6.65 \times 10^{-27} \text{ kg}$$

De acuerdo con la ecuación de de Broglie, la longitud de onda es:

$$\lambda = \frac{h}{p} = \frac{h}{mv} = \frac{6.63 \times 10^{-34}\ \text{joule-seg}}{6.65 \times 10^{-27}\ \text{kg} \times 1.635 \times 10^{3}\ \text{m/seg}} = 0.609 \times 10^{-10}\ \text{m}$$

$$= 0.609 \times 10^{-8}\ \text{cm}$$

Este resultado es 1.5% mayor que el valor medida por difracción en un cristal, pero está dentro de las límites de error del experimenta. ◀

Experimentos como el considerada en el ejemplo 3-2 son muy difíciles de realizar, ya que las intensidades que se obtienen en haces atómicas son muy pequeñas. Los experimentos de difracción de neutranes, utilizando cristales con distancia interatómica conocida, proporcionan una confirmación precisa de la ecuación de de Broglie y de la existencia de ondas de materia. La precisión se debe a que la fuente de neutrones en reactores nucleares es muy abundante y, efectivamente, la difracción de neutrones es hoy día un método importante para el estudio de estructuras cristalinas. Ciertos cristales, como los orgánicos hidrogenados, son particularmente adecuados para análisis con difracción de neutrones, ya que los neutrones son fuertemente dispersados por los átomos de hidrógeno, mientras que los rayos x son dispersados muy débilmente por dichos átomos. Principalmente, los rayos X interactúan con los electrones en el átomo y los electrones interactúan con la carga nuclear así como con los electrones atómicos por fuerzas electromagnéticas, de modo que su interacción con átomos de hidrógeno es débil, ya que su carga es pequeña. Los neutrones, sin embargo, interactúan con el núcleo de los átomos por medio de fuerzas nucleares y la interacción es fuerte.

3.2 La dualidad onda-partícula

En la física clásica, la energía es transportada por partículas o por ondas. El físico clásico observaba las ondas de agua transportando energía sobre la superficie del agua o balas transportando energía del cañón al blanco. De tales experiencias, construyeron un modelo ondulatorio de ciertos fenómenos macroscópicos y un modelo corpuscular para otros fenómenos macroscópicos, extrapolando naturalmente estos modelos a regiones menos accesibles al ojo humano. Así por ejemplo, se explicó la propagación del sonido en base al modelo ondulatorio y la presión de los gases en base al modelo corpuscular (teoría cinética). El éxito obtenido los condicionó a esperar que todo ente debiera ser o partícula u onda. De hecho, estos éxitos llegaron a los primeros años del siglo veinte, con las aplicaciones de la teoría ondulatoria de Maxwell a la radiación y el descubrimiento de partículas elementales de materia, tales como el neutrón y el positrón.

Así pues, el físico clásico no estaba preparado para descubrir que para poder entender la radiación, en algunas situaciones era necesario invocar un modelo corpuscular, como en el efecto Compton, mientras que en otras situaciones, un modelo ondulatorio, como en la difracción de rayos X. Pero quizás, sea más sorprendente el hecho de que esta dualidad onda-partícula se aplique, tanto a la materia, como a la radiación. La relación carga-masa del electrón y su rastro de ionización en la materia (secuencia de colisiones localizadas) sugieren un modelo corpuscular, sin embargo, la difracción de electrones sugiere un modelo corpuscular. El físico actual sabe que, a un ente dado, debe aplicarle ambos modelos. Sin embargo, es importante hacer notar que, en una medición dada, se debe aplicar un solo modelo, ya que no se puede utilizar ambos modelos bajo las mismas circunstancias. Cuando una partícula es detectada, mediante algún tipo de interacción, actúa como partícula, en el sentido de que está localizada; cuando está en movimiento, actúa como onda, en el sentido de que se observan fenómenos de interferencia y, desde luego, una onda se extiende y no está localizada.

Neils Bohr resumió esta situación en su *principio de complementaridad*. Los modelos corpuscular y ondulatorio son complementarios; si una medida prueba el carácter ondulatorio de la radiación o la materia, entonces es imposible probar la naturaleza corpuscular en el mismo

xperimento y viceversa. El modelo que se utilice lo determina la naturaleza del experimento. demás, el conocimiento de la radiación o la materia será incompleto, a menos que se nsideren medidas que revelen tanto los aspectos ondulatorios como los corpusculares. Así ues, la radiación y la materia no son ni simplemente ondas ni simplemente partículas. Para escribir este comportamiento, se requiere un modelo más general, y desde el punto de vista ásico, más complicado, a pesar de que en casos extremos sea aplicable un modelo ondulatorio mple o un modelo corpuscular simple.

La conexión entre los modelos corpuscular y ondulatorio, se encuentra en una interpretación robabilística de la dualidad onda-partícula. En el caso de la radiación, fue Einstein quien nificó las teorías ondulatoria y corpuscular; subsecuentemente Max Born aplicó un argumento milar para unificar las teorías ondulatoria y corpuscular de la materia.

Desde el punto de vista ondulatorio, la intensidad de radiación, I, es proporcional a $\overline{\mathscr{E}^2}$, nde $\overline{\mathscr{E}^2}$ es el promedio sobre un ciclo del cuadrado de la intensidad de campo eléctrico de la da. (I es el valor promedio del llamado vector de Poynting y se utiliza el símbolo $\mathscr{E}$ en lugar E, para el campo eléctrico, para evitar se confunda con la energía total E). Desde el punto de sta de los fotones o corpuscular, la intensidad de la radiación se escribe como $I=Nh\nu$, donde N el número promedio de fotones que cruzan una unidad de área, perpendicular a la dirección propagación, por unidad de tiempo. Fue Einstein quien sugirió que $\overline{\mathscr{E}^2}$, que en la teoría ectromagnética es proporcional a la energía radiante en una unidad de volumen, podría terpretarse como una medida del número promedio de fotones por unidad de volumen.

Recuérdese que Einstein introdujo la idea de granularidad en la radiación, abandonando la terpretación continua de Maxwell. Esto conduce a una interpretación estadística de la tensidad. Desde este punto de vista, una fuente puntual emite fotones al azar en todas recciones. El número promedio de fotones que crucen una unidad de área, decrecerá a medida e aumente la distancia entre la fuente y el área. Esto se debe a que los fotones se dispersan bre una esfera cuya área es mayor a medida que están más alejados de la fuente. Como el área una esfera es proporcional al cuadrado de su radio, en promedio, se obtiene una ley versamente proporcional al cuadrado para la intensidad, de igual modo que en el caso dulatoria. Desde el punto de vista ondulatorio, puede imaginarse que, desde una fuente, se neran ondas esféricas cuya intensidad disminuye en forma inversamente proporcional a la stancia de la fuente. En este caso, estas ondas se pueden considerar como ondas guía para los tones, cuya intensidad puede medirse por $\overline{\mathscr{E}^2}$, las ondas por sí mismas, no poseen energía —solo n fotones— pero son una construcción artificial cuya intensidad mide el número promedio fotones por unidad de volumen.

Se ha utilizado la palabra "promedio", porque los procesos de emisión son de naturaleza tadística. No se especifica cuántos fotones cruzan una unidad de área por unidad de tiempo, lo su número promedio; su número exacto puede fluctuar en el espacio y en el tiempo, del smo modo como en la teoría cinética de los gases existen fluctuaciones de muchas cantidades specto a un valor promedio. Sin embargo, se puede afirmar categóricamente, que la obabilidad de que un fotón cruce una unidad de área a 3 m de la fuente, es exactamente un veno de la probabilidad de que un fotón cruce una unidad de área a 1 m de la fuente. Por lo nto, en la fórmula $I=Nh\nu$, N es un valor promedio que es una medida de la probabilidad de contrar a un fotón cruzando una unidad de área en una unidad de tiempo. Si se igualan las presiones ondulatoria y corpuscular, se tiene;

$$I = (1/\mu_0 c)\overline{\mathscr{E}^2} = h\nu N$$

modo que $\overline{\mathscr{E}^2}$ es proporcional a N. Entonces, la interpretación de Einstein de $\overline{\mathscr{E}^2}$, como edida probabilística de la densidad de fotones, se va aclarando. Se espera que, como en la teoría nética, las fluctuaciones respecto a un promedio sean más notorias a bajas intensidades que a as intensidades, de modo que, en ese caso, los fenómenos granulares cuánticos contradicen s dramáticamente el punto de vista del espectro continuo clásico.

En analogía al punto de vista de Einstein de la radiación, Max Born propuso una unión similar de la dualidad onda-partícula de la materia. Esto sucedió varios años después de que Schrödinger desarrolló su generalización del postulado de de Broglie, llamada mecánica cuántica. La teoría de Schrödinger será examinada cuantitativamente en capítulos posteriores. Aquí sólo se tratará de usar la idea de Born, en forma cualitativa para conceptualmente preparar el escenario para un análisis detallado subsecuente.

Se asociará a las ondas de materia algo más que una longitud de onda y una frecuencia. Esto se hará introduciendo una función que represente a la onda de de Broglie, llamada *función de onda*. Por ejemplo, para partículas moviéndose en la dirección x con un valor preciso de su impulso lineal y su energía, la función de onda se puede escribir simplemente como una función senoidal de amplitud A, tal como:

$$\Psi(x,t) = A \operatorname{sen} 2\pi\left(\frac{x}{\lambda} - \nu t\right) \tag{3-4a}$$

Esto es análogo a

$$\mathscr{E}(x,t) = A \operatorname{sen} 2\pi\left(\frac{x}{\lambda} - \nu t\right) \tag{3-4b}$$

para el campo eléctrico de una onda electromagnética senoidal con longitud de onda λ, y frecuencia ν, moviéndose en la dirección x. La cantidad $\overline{\Psi^2}$ para ondas de materia, se comportará de manera análoga al de $\overline{\mathscr{E}^2}$ para ondas de radiación. Esa cantidad, el promedio del cuadrado de la función de onda de ondas de materia, es una medida de la probabilidad de encontrar a una partícula por unidad de volumen, en determinado lugar a un tiempo dado. Así como $\mathscr{E}$ es una función del tiempo y del espacio, también lo es Ψ, y como se verá después, así como $\mathscr{E}$ satisface una ecuación de onda, así también Ψ (la ecuación de Schrödinger). La cantidad $\mathscr{E}$ es una onda (de radiación) asociada con un fotón, mientras que Ψ es una onda (de materia) asociada con una partícula.

Como dice Born: "De acuerdo con este punto de vista, el curso completo de eventos está gobernado por las leyes de la probabilidad, a un estado en el espacio, le corresponde una probabilidad definitiva, la cual es dada por la onda de de Broglie asociada con el estado. Por lo tanto, un proceso mecánico es acompañado por un proceso ondulatorio, la onda-guía, descrita por la ecuación de Schrödinger, cuya importancia está en dar la probabilidad de un curso definitivo del proceso mecánico. Si por ejemplo, la amplitud de la onda-guía es cero en cierto punto del espacio, esto quiere decir que la probabilidad de encontrar al electrón en ese punto es despreciablemente pequeña".

Así como desde el punto de vista de Einstein de la radiación, no se especifica la posición exacta de un fotón, sino que en su lugar se especifica $\overline{\mathscr{E}^2}$, la probabilidad de encontrar un fotón en cierta posición a un tiempo dado; así también, desde el punto de vista de Born, no se especifica la posición exacta de una partícula a un tiempo determinado, sino que se especifica $\overline{\Psi^2}$, la probabilidad de encontrar a la partícula en cierta posición a un tiempo dado. Además, así como se acostumbra a sumar funciones de onda ($\mathscr{E}_1 + \mathscr{E}_2 = \mathscr{E}$) para dos ondas electromagnéticas superpuestas, cuya intensidad resultante se obtiene de $\mathscr{E}^2$, así también se podían sumar las funciones de onda de dos ondas de materia superpuestas ($\Psi_1 + \Psi_2 = \Psi$) cuya intensidad resultante será dada por $\overline{\Psi^2}$. Es decir, un *principio de superposición* es aplicable tanto a ondas de materia como a radiación. Esto concuerda con el sorprendente hecho experimental, de que la materia presenta propiedades de interferencia y difracción, hecho que simplemente no puede explicarse en base a las ideas de la mecánica clásica. Dos ondas siempre se pueden combinar de modo tal que resulte una onda de intensidad grande o que se cancelen, ya que pueden superponerse constructivamente (en fase) o destructivamente (fuera de fase), pero dos partículas clásicas no se pueden combinar de modo tal que se cancelen.

Posiblemente, el estudiante aceptará lo lógico de la fusión de los conceptos de onda y de partícula y, sin embargo, se preguntará si una interpretación probabilística o estadística es necesaria. En 1927, Heisenberg y Bohr fueron los primeros en demostrar que en la unión de las descripciones ondulatoria y corpuscular de la materia y la radiación, es esencial el concepto de probabilidad. Estos temas serán investigados en las siguientes secciones.

3.3 El principio de incertidumbre

El uso de consideraciones probabilísticas, no es ajeno a la física clásica. Por ejemplo, la mecánica estadística clásica utiliza la teoría de la probabilidad. Sin embargo, las leyes básicas de la física clásica (tales como las leyes de Newton) son deterministas, y el análisis estadístico es simplemente un instrumento que se utiliza en el tratamiento de sistemas muy complicados. Por otro lado, de acuerdo con Heisenberg y Bohr, el punto de vista probabilístico es fundamental en la física cuántica y el determinismo debe ser desartado. A continuación se verá cómo se llega a esta conclusión.

En mecánica clásica, pueden resolverse las ecuaciones de movimiento de un sistema con fuerzas dadas, para obtener la posición y el impulso de una partícula para todo valor del tiempo. Todo lo que se necesita conocer es la posición e impulso de la partícula para algún valor del tiempo, $t = 0$ (las condiciones iniciales), y el movimiento futuro será determinado exactamente. Esta mecánica ha sido utilizada con gran éxito en el mundo macroscópico, en astronomía por ejemplo, para predecir los movimientos subsecuentes de objetos en base a sus movimientos iniciales. Sin embargo, nótese, que en el proceso de hacer observaciones, el observador interactúa con el sistema. Un ejemplo de la astronomía moderna, es la medición precisa de la posición de la luna, haciendo rebotar ondas de radar en ella. El movimiento de la luna es perturbado por la medición, pero su masa es tan grande que la perturbación se puede ignorar. En una escala más pequeña, como un experimento macroscópico bien diseñado sobre la tierra, dichas perturbaciones son también generalmente muy pequeñas o al menos, controlables y se pueden tomar en cuenta de antemano por cálculos adecuados. Por lo tanto, el físico clásico suponía, de modo natural, que en el campo de los sistemas microscópicos, la posición y el impulso de un objeto, como un electrón, podrían determinarse exactamente por medio de observaciones similares. Heisenberg y Bohr pusieron en tela de juicio esta suposición.

La situación es muy similar a la que existía al nacimiento de la teoría de la relatividad. Los físicos hablaban de intervalos de tiempo y longitud, es decir, espacio y tiempo, sin preguntarse críticamente cómo son medidos en realidad. Por ejemplo, se hablaba de la simultaneidad de dos eventos separados, sin siquiera preguntarse cómo se podría establecer dicha simultaneidad. Einstein demostró que la simultaneidad no era un concepto absoluto, como se había supuesto antes, sino que dos eventos que son simultáneos para un observador, ocurren a tiempos diferentes para otro observador que se mueve respecto al primero. La simultaneidad es un concepto relativo. Así pues, de modo similar, debe preguntarse cómo se mide realmente la posición y el impulso.

En un experimento real, ¿se puede determinar la posición y el impulso de la materia o la radiación, en el mismo instante? La respuesta que se obtiene de la teoría cuántica, es: no más exactamente que lo permitido por el *principio de incertidumbre* de Heisenberg. Existen dos partes de este principio, también llamado principio de indeterminación. La primera parte implica la medición simultánea de la posición y el impulso. Afirma que en un experimento no se puede determinar simultáneamente el valor exacto de una componente del impulso, es decir, p_x de una partícula y también el valor exacto de la coordenada correspondiente x. Por otro lado, la precisión en la medición estará inherentemente limitada por el proceso de medida en sí, de modo tal que

donde el impulso p_x se conoce hasta una incertidumbre de Δp_x y al mismo tiempo, la posición x hasta una incertidumbre de Δx. Aquí, $\hbar$ (Léase h-barra) es una notación simplificada en lugar de $h/2\pi$, donde h es la constante de Planck. Es decir;

$$\hbar \equiv h/2\pi$$

Existen las relaciones correspondientes para las otras componentes del impulso, a saber, $\Delta p_y \Delta y \geq \hbar/2$ y $\Delta p_z \Delta z \geq \hbar/2$; y asimismo para el impulso angular. Es importante hacer notar que este principio nada tiene que ver con los adelantos en la instrumentación que conduzcan a mejores determinaciones simultáneas de p_x y x. Lo que el principio afirma es que aún con instrumentos ideales, nunca se puede superar $\Delta p_x \Delta x \geq \hbar/2$. Obsérvese también que se implica un *producto* de incertidumbres, de modo que, por ejemplo, entre más se modifique un experimento para mejorar p_x, más se sacrifica la habilidad de poder determinar x con precisión. Si p_x se conociera exactamente, se ignoraría totalmente la x (es decir, si $\Delta p_x = 0$, $\Delta x = \infty$). Por lo tanto, la *restricción no está en la precisión con la que se pueda medir* x *o* p_x, *sino en el producto* $\Delta p_x \Delta x$ en una medida simultánea de ambas.

La segunda parte del principio de incertidumbre está relacionada con medidas de la energía E, y el tiempo t necesarios para medir, por ejemplo, el intervalo Δt durante el cual un fotón es emitido de un átomo con una variación en su energía ΔE. En este caso

$$\Delta E \Delta t \geq \hbar/2 \qquad (3\text{-}6)$$

donde ΔE es la incertidumbre en el conocimiento de la energía E del sistema y Δt el intervalo característico de la rapidez de cambio del sistema.

Más adelante se demostrará que las relaciones de Heisenberg se pueden derivar del postulado de de Broglie y de propiedades sencillas comunes a toda onda. Dado que el postulado de de Broglie se puede verificar por los experimentos descritos anteriormente, se puede afirmar que el principio de incertidumbre se basa en el experimento. Asimismo, pronto se verá que este principio es consistente con otros experimentos. Sin embargo, primero se debe hacer notar que, una vez más, la constante de Planck sirve para distinguir los resultados cuánticos de los clásicos. Si h, o $\hbar$, fuera cero en (3-5) y (3-6), no habría ninguna limitación básica en las medidas, como es el caso en el punto de vista clásico. Una vez más, la pequeñez de h, es lo que pone al principio fuera de las experiencias ordinarias. Lo anterior es análogo a la pequeñez de la relación v/c en situaciones macroscópicas, que pone a la relatividad fuera de alcance en experiencias ordinarias. Por lo tanto, en principio, la física clásica tiene una validez limitada y, en el dominio microscópico, conducirá a contradicciones con los resultados experimentales, ya que si x y p no pueden determinarse simultáneamente, entonces las condiciones iniciales del movimiento no se pueden especificar exactamente y consecuentemente no se podrá determinar con precisión el comportamiento futuro de un sistema. En lugar de hacer predicciones deterministas sólo se pueden enunciar los resultados más probables de una observación, dando las probabilidades relativas de que ocurran. En efecto, como el hecho de observar un sistema lo perturba de modo no completamente predecible, la observación cambia el movimiento anterior del sistema a un estado de movimiento nuevo que no puede ser conocido completamente.

En lo siguiente se mostrará el origen físico del principio de incertidumbre, y en este proceso se obtendrá una visión más clara que ayudará a apreciar mejor una demostración más formal que se da en la sección subsecuente. Primeramente, se utiliza un experimento de pizarrón debido a Bohr, para verificar (3-5). Se propone medir, tan exactamente como sea posible, la posición de una partícula "puntual", como un electrón. Para mayor precisión, se utiliza un microscopio para observar al electrón, como el que se ilustra en la figura 3-6. Para poder ver al electrón, es necesario iluminarlo, ya que lo que ve el observador, son los fotones de la luz dispersados por el electrón. En este punto, aún sin hacer cálculo alguno, se puede ver cómo

aparece el principio de incertidumbre. El propio acto de observar al electrón lo perturba. En el momento de iluminar al electrón, éste retrocede a consecuencia del efecto Compton, de modo tal que, como se verá, no se puede determinar completamente. Sin embargo, si el electrón no está iluminado, no puede ser observado (detectado). Por lo tanto, el principio de incertidumbre se refiere al proceso mismo de medición y expresa el hecho de que siempre existe una interacción indeterminada entre el observador y lo observado; no existe nada que pueda hacerse para evitar esta interacción o para compensarla por adelantado. En el caso que se está considerando, se puede tratar de reducir la perturbación al electrón lo más posible, utilizando una fuente de luz débil. La fuente más débil que se puede obtener será tal que se observará al electrón sólo si *un solo fotón* dispersado entra en la lente objetivo. La magnitud del impulso del

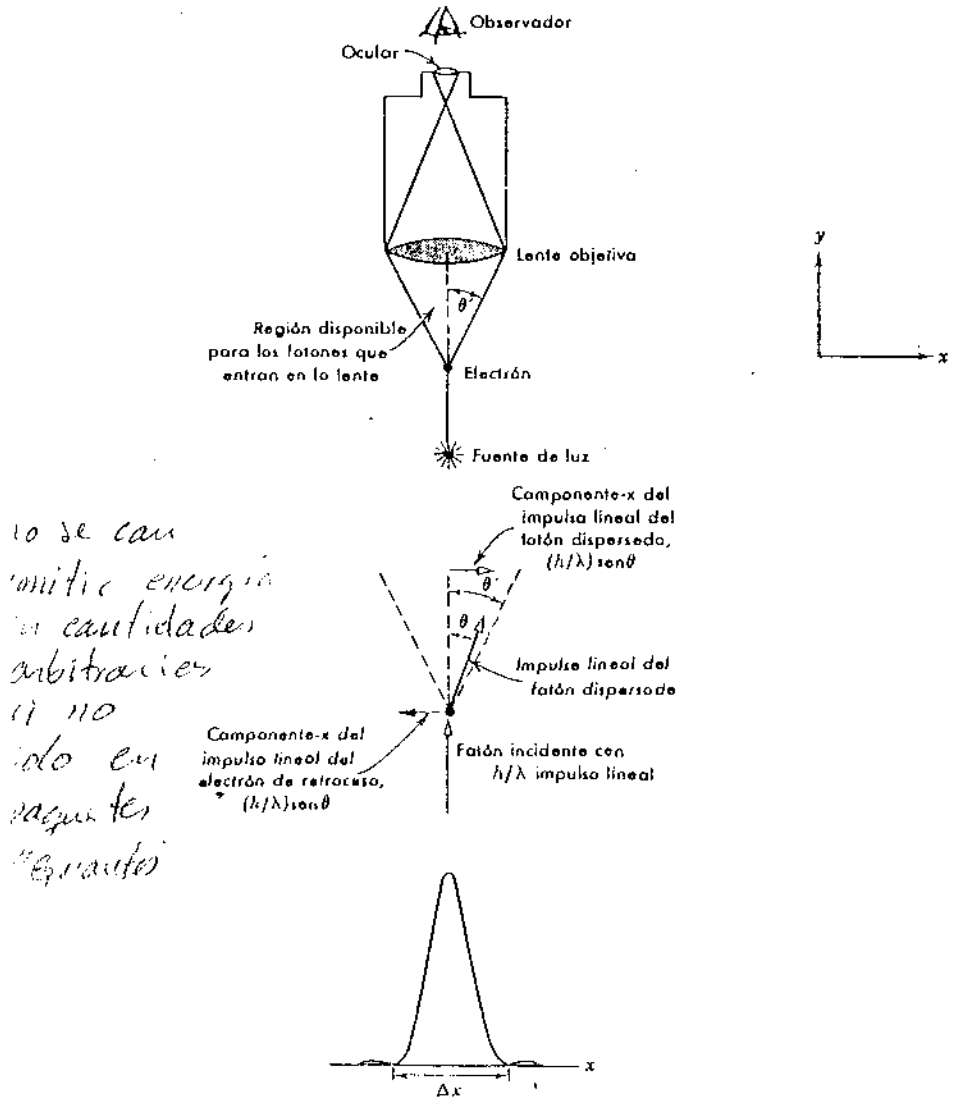


FIGURA 3-6

Microscopio del experimento de pizarrón de Bohr. *Arriba:* el aparato. *Centro:* dispersión de un fotón de iluminación por un electrón. *Abajo:* imagen del patrón de difracción del electrón, vista por el observador.

fotón es $p = h/\lambda$. Pero el fotón puede ser dispersado en *cualquier dirección* dentro del ángulo $2\theta'$ que subtiende la lente objetivo en el electrón. Por esta razón, no es posible considerar la interacción. Por lo tanto, se encuentra que la componente x del impulso del fotón puede variar desde $+p$ sen θ' hasta $-p$ sen θ' y su incertidumbre después de la interacción está dada por:

$$\Delta p_x = 2p \operatorname{sen} \theta' = (2h/\lambda) \operatorname{sen} \theta'$$

Por la conservación del impulso, se requiere que el electrón reciba impulso de retroceso en la dirección x que será igual en magnitud al cambio en la magnitud de la componente x del impulso del fotón y, por lo tanto, la componente x del impulso del electrón, tendrá una incertidumbre dada por la misma cantidad. Obsérvese que para reducir Δp_x, se puede utilizar luz de mayor longitud de onda o utilizar un microscopio cuya lente objetivo subtienda un ángulo menor.

¿Qué hay de la posición del electrón a lo largo de x? Recuérdese que la imagen de un punto, en un microscopio, no es un punto, sino un patrón de difracción; la imagen del electrón es "borrosa". El poder de resolución del microscopio determina la exactitud última a la cual se puede localizar el electrón. Si como medida de la incertidumbre en x se toma la amplitud del máximo de difracción central, una expresión bien conocida para el poder de resolución de un microscopio da:

$$\Delta x = \lambda/\operatorname{sen}\theta'$$

(nótese que como sen $\theta \simeq \theta$, este es un ejemplo de la relación general, $a \simeq \lambda/\theta$ entre la dimensión característica de un aparato de difracción, la longitud de onda de las ondas difractadas y el ángulo de difracción). Entonces, el único fotón disponible debió originarse en *algún* lugar dentro de esta región del eje del microscopio, de manera tal que la incertidumbre en la posición es Δx. (No se puede asegurar exactamente en dónde se origina un fotón dado, a pesar de que un número grande de repeticiones del experimento, los fotones que forman la imagen total, producirán el patrón de difracción que se muestra en la figura). Obsérvese que para reducir Δx se puede analizar luz de longitud de onda más corta o una lente objetivo que subtienda un ángulo mayor.

Tomando el producto de las incertidumbres, se encuentra

$$\Delta p_x \Delta x = \left(\frac{2h}{\lambda} \operatorname{sen} \theta'\right)\left(\frac{\lambda}{\operatorname{sen} \theta'}\right) = 2h \qquad (3\text{-}7)$$

en concordancia razonable con el límite último $h/2$ que fija el principio de incertidumbre. No es posible hacer Δp_x y Δx *simultáneamente*, tan pequeñas como se quiera, ya que el procedimiento que reduce una de ellas incrementa la otra. Por ejemplo, si se utiliza luz de longitud de onda corta (por ejemplo rayos γ) para reducir Δx obteniendo una mejor resolución, se aumenta el retroceso de Compton y aumenta Δp_x, e inversamente. De hecho, la longitud de onda, λ y el ángulo θ' subtendido por la lente objetivo, ni siquiera aparecen en el resultado. En la práctica, los resultados experimentales serán peores que lo que se sugiere en (3-7), ya que dicho resultado representa el caso más ideal posible. Sin embargo, se llega a ese resultado a partir de fenómenos físicos genuinamente mensurables, a saber, el efecto Compton y el poder de resolución de una lente.

En realidad, el estudiante no debe pensar que hay algo de misterioso en este resultado. Es un resultado directo de la cuantización de la radiación. Fue necesario tener al menos un fotón iluminando al electrón, o de otra forma no habría iluminación; y aún un solo fotón lleva un impulso de magnitud $p = h/\lambda$. Es este único fotón el que provee la interacción entre el microscopio y el electrón. Esta interacción perturba la partícula de modo tal, que no se puede predecir ni controlar y el resultado es que ni las coordenadas ni el impulso pueden conocerse completamente después de una medición. Si las leyes de la física clásica fuesen válidas, entonces

la radiación sería considerada como continua en lugar de granular y la iluminación se podría reducir a niveles arbitrariamente pequeños al mismo tiempo que se utilizarían longitudes de onda arbitrarias para obtener una resolución "perfecta". En principio no habría un límite inferior simultáneamente para la resolución y el impulso de retroceso y no habría principio de incertidumbre. Pero esto no se puede hacer, un fotón es indivisible. Una vez más se puede ver de $\Delta p_x \Delta x \geq h/2$, que la constante de Planck es una medida de la mínima perturbación incontrolable que distingue a la física cuántica de la física clásica.

Ahora se considerará (3-6) que relaciona la incertidumbre en la energía y en el tiempo. Para el caso de una partícula libre, (3-6) puede obtenerse de (3-5), que relaciona la posición y el impulso en la forma siguiente. Considérese un electrón que se mueve a lo largo del eje x y cuya energía se puede escribir como $E = p_x^2/2m$. Si la incertidumbre en p_x está dada por Δp_x, entonces la incertidumbre en E estará dada por $\Delta E = (p_x/m)\Delta p_x = v_x \Delta p_x$. En este caso v_x puede interpretarse como la velocidad de retroceso a lo largo de x, del electrón que es iluminado para medir su posición o si el tiempo requerido para una medida es Δt, entonces, la incertidumbre en su posición x es $\Delta x = v_x \Delta t$. Combinando $\Delta t = \Delta x/v_x$ y $\Delta E = v_x \Delta p_x$, se obtiene $\Delta E \Delta t = \Delta p_x \Delta x$. Pero $\Delta p_x \Delta x \geq h/2$. Por lo tanto:

$$\Delta E \Delta t \geq h/2$$

Ejemplo 3-3. La velocidad de una bala ($m = 50$g) y la velocidad de un electrón, ($m = 9.1 \times 10^{-28}$ g) se miden y resultan iguales, a 300 m/seg, con una incertidumbre de 0.01%. ¿Cuál será la exactitud fundamental con que se podrá determinar la posición de cada una, si la posición y la velocidad se miden simultáneamente en el mismo experimento?

Para el electrón

$$p = mv = 9.1 \times 10^{-31} \text{ kg} \times 300 \text{ m/seg} = 2.7 \times 10^{-28} \text{ kg-m/seg}$$

y

$$\Delta p = m\Delta v = 0.0001 \times 2.7 \times 10^{-28} \text{ kg-m/seg} = 2.7 \times 10^{-32} \text{ kg-m/seg}$$

de modo que;

$$\Delta x \geq \frac{h}{4\pi \Delta p} = \frac{6.6 \times 10^{-34} \text{ joule-seg}}{4\pi \times 2.7 \times 10^{-32} \text{ kg-m/seg}} = 2 \times 10^{-3} \text{ m} = 0.2 \text{ cm}$$

Para la bala

$$p = mv = 0.05 \text{ kg} \times 300 \text{ m/seg} = 15 \text{ kg-m/seg}$$

$$\Delta p = 0.0001 \times 15 \text{ kg-m/seg} = 1.5 \times 10^{-3} \text{ kg-m/seg}$$

de modo que;

$$\Delta x \geq \frac{h}{4\pi \Delta p} = \frac{6.6 \times 10^{-34} \text{ joule-seg}}{4\pi \times 1.5 \times 10^{-3} \text{ kg-m/seg}} = 3 \times 10^{-32} \text{ m.}$$

Por lo tanto, para objetos macroscópicos, tales como balas, el principio de incertidumbre no impone límites prácticos en el procedimiento de medición, ya que Δx en este caso resulta aproximadamente 10^{-17} veces el diámetro de un núcleo; pero para objetos microscópicos, como electrones, existen límites prácticos, ya que en este ejemplo Δx resulta 10^7 veces el diámetro de un átomo. ◀

3.4 Propiedades de ondas de materia

En esta sección se derivará el principio de incertidumbre combinando las relaciones de de Broglie-Einstein, $p = h/\lambda$ y $E = h\nu$, con propiedades matemáticas simples que son univer-

sales a todas las ondas. Se empezará el desarrollo de estas propiedades llamando la atención a una paradoja aparente.

La velocidad de propagación w para una onda con longitud de onda λ y frecuencia ν, está dada por la relación familiar, que más adelante se verificará

$$w = \lambda\nu \tag{3-8}$$

Evaluando w para una onda de de Braglie asociada con una partícula de impulso p y energía total E se obtiene;

$$w = \lambda\nu = \frac{h}{p}\frac{E}{h} = \frac{E}{p}$$

En seguida se supone que la partícula se mueve a una velocidad v no relativista en una región de energía potencial cero. (La validez de las conclusiones no será limitada por estas suposiciones). Evaluando p y E en términos de v y de la masa m de la partícula se encuentra

$$w = \frac{E}{p} = \frac{mv^2/2}{mv} = \frac{v}{2} \tag{3-9}$$

Este resultado es, en apariencia preocupante, pues parecía que la onda de materia no viaja a la misma velocidad que la partícula que representa, pero en realidad no existe dificultad alguna como se demostrará en la discusión que sigue.

Imagínese a la partícula moviéndose a lo largo del eje x, sin la acción de fuerzas ya que su energía potencial tiene el valor constante cero. También moviéndose a lo largo del eje x se tiene a su onda de materia asociada. Supóngase también, en este experimento de pizarrón, que a lo largo del eje x se han distribuido un conjunto (hipotético) de instrumentos capaces de medir la amplitud de la onda de materia. Para cierto tiempo $t = 0$, se registran las lecturas en estos instrumentos. Los resultados del experimento pueden presentarse en una gráfica de los valores instantáneos de la onda, que se designan por $\Psi(x, t)$, como función de la distancia x, al tiempo fijo $t = 0$. En este momento no se necesita saber mucho acerca de ondas de materia para darse cuenta que *cualitativamente* la gráfica debe asemejarse a la que se muestra en la figura 3-7. La amplitud de la onda de materia, deberá estar modulada de modo tal, que su valor será diferente de cero sólo en una región finita del espacio en la vecindad de la partícula. Esto es necesario porque la onda de materia, de algún modo, debe estar espacialmente asociada con la partícula cuyo movimiento controla. La onda de materia tendrá la forma de un grupo de ondas y a medida que transcurra el tiempo, el grupo seguramente deberá moverse a lo largo del eje x con la misma velocidad que la partícula.

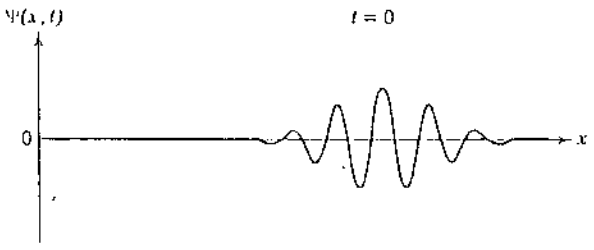


FIGURA 3-7

Una onda de de Broglie para una partícula.

El estudiante podrá recordar, de su estudio del movimiento de ondas clásicas, que para el movimiento de un grupo de ondas es necesario distinguir entre la velocidad g del grupo y la velocidad w, muy diferente, de las oscilaciones de las ondas individuales. Desde luego esto resulta alentador, pero debe demostrarse que g es igual a la velocidad de la partícula. Para esto, se desarrollará una relación entre g y las cantidades ν y λ, comparable con la relación (3-8) entre w y estas cantidades.

Se empezará considerando el caso más simple de movimiento ondulatorio, una onda senoidal de frecuencia ν y longitud de onda λ, que tiene amplitud constante y unitaria desde $-\infty$ hasta $+\infty$, pero que se mueve con velocidad uniforme en la dirección en que x aumenta. Una onda tal, puede representarse matemáticamente por la función:

$$\Psi(x,t) = \operatorname{sen} 2\pi\left(\frac{x}{\lambda} - \nu t\right) \tag{3-10a}$$

o en una forma más conveniente

$$\Psi(x,t) = \operatorname{sen} 2\pi(\kappa x - \nu t) \qquad \text{donde } \kappa \equiv 1/\lambda \tag{3-10b}$$

Que lo anterior representa la onda descrita, puede verse de las consideraciones siguientes:

1. Fijando la x, en cualquier valor, se observa que la función oscila senoidalmente con frecuencia ν y amplitud uno.

2. Fijando t, se observa que la función tiene una dependencia senoidal con x, con longitud de onda λ o *longitud de onda recíproca* κ.

3. Los ceros de la función, que corresponden a los nodos de la onda que representa, se encuentran en las posiciones x_n para las cuales

$$2\pi(\kappa x_n - \nu t) = \pi n \qquad\qquad n = 0, \pm 1, \pm 2, \ldots$$

o

$$x_n = \frac{n}{2\kappa} + \frac{\nu}{\kappa} t$$

Por lo tanto, estos nodos y de hecho todos los puntos sobre la onda se mueven en la dirección en que x crece, con velocidad

$$w = dx_n/dt$$

que es igual a

$$w = \nu/\kappa$$

Nótese que lo anterior es idéntico a (3-8), ya que $\kappa = 1/\lambda$.

En seguida se analizará el caso en que se modula la amplitud de las ondas para formar un grupo. Matemáticamente se puede obtener un grupo de ondas moviéndose en la dirección en que x crece, similar al grupo de ondas de materia que se ilustra en la figura 3-7, sumando un número infinitamente grande de ondas de la forma (3-10b), cada una con frecuencia ν y longitud de onda recíproca κ infinitamente diferentes. (Pronto se explicará cómo sucede lo anterior). Sin embargo, las técnicas matemáticas se complican y para el propósito presente, será suficiente considerar qué sucede si se suman sólo dos de esas ondas. Por lo tanto, se toma

$$\Psi(x,t) = \Psi_1(x,t) + \Psi_2(x,t) \tag{3-11}$$

donde

$$\Psi_1(x,t) = \operatorname{sen} 2\pi[\kappa x - \nu t]$$

y

$$\Psi_2(x,t) = \operatorname{sen} 2\pi[(\kappa + d\kappa)x - (\nu + d\nu)t]$$

Ahora bien,

$$\operatorname{sen} A + \operatorname{sen} B = 2\cos[(A - B)/2]\operatorname{sen}[(A + B)/2]$$

Aplicando lo anterior al caso que nos ocupa, se tiene

$$\Psi(x,t) = 2\cos 2\pi\left[\frac{d\kappa}{2}x - \frac{d\nu}{2}t\right]\operatorname{sen} 2\pi\left[\frac{(2\kappa + d\kappa)}{2}x - \frac{(2\nu + d\nu)}{2}t\right]$$

como $d\nu \ll 2\nu$ y $d\kappa \ll 2\kappa$, esto es:

$$\Psi(x,t) = 2\cos 2\pi\left(\frac{d\kappa}{2}x - \frac{d\nu}{2}t\right)\operatorname{sen} 2\pi(\kappa x - \nu t) \tag{3-12}$$

En la figura 3-8 se muestra una gráfica de $\Psi(x, t)$ como función de x para el valor fijo $t = 0$. El segundo término de $\Psi(x, t)$ es una onda de la misma forma que (3-10b), pero modulada por el primer término, de modo que las oscilaciones de $\Psi(x, t)$ caen dentro de la envolvente cuya amplitud varía periódicamente. Dos ondas cuyas longitudes de onda recíproca y frecuencia difieren ligeramente, se interfieren y se refuerzan alternativamente de modo tal que se producen grupos sucesivos. Estos grupos y las ondas individuales que contienen, se mueven en la dirección en que x crece. La velocidad w de las ondas individuales se pueden evaluar considerando el segundo término de $\Psi(x, t)$ y la velocidad g de los grupos se puede evaluar del primer término. Una vez más, procediendo como en la consideración 3, se obtiene

$$w = \frac{\nu}{\kappa} \tag{3-13a}$$

y también el resultado nuevo

$$g = \frac{d\nu/2}{d\kappa/2} = \frac{d\nu}{d\kappa} \tag{3-13b}$$

Para un número infinito de ondas que se combinen para formar un grupo en movimiento, se puede demostrar que la dependencia de la *velocidad de onda* w y la *velocidad de grupo* g con ν, κ y

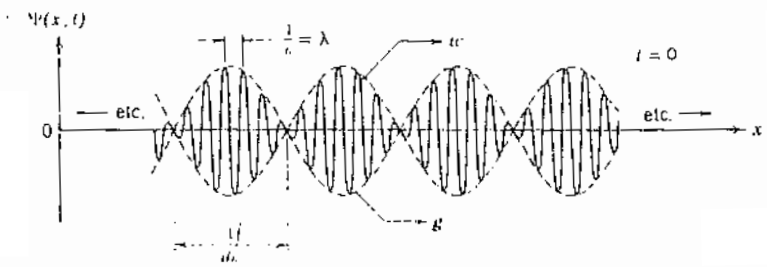


FIGURA 3-8

Suma de dos ondas senoidales de frecuencias y longitudes de onda recíproca κ, ligeramente diferentes.

$d\nu/d\kappa$ es exactamente la misma que para el caso sencillo que se ha considerado. Las ecuaciones (3-13a) y (3-13b) tienen validez general.

Finalmente, se tienen las condiciones para poder calcular la velocidad de grupo g del grupo de ondas asociada con la partícula en movimiento. De las relaciones de Einstein y de de Broglie se tiene

$$\nu = E/h \qquad \text{y} \qquad \kappa \equiv 1/\lambda = p/h$$

de modo que

$$d\nu = dE/h \qquad \text{y} \qquad d\kappa = dp/h$$

Por lo tanto, la velocidad de grupo es:

$$g = d\nu/d\kappa = dE/dp$$

Igualando

$$E = \frac{mv^2}{2} \qquad \text{y} \qquad p = mv$$

se obtiene

$$\frac{dE}{dp} = \frac{mv\,dv}{m\,dv} = v$$

de donde se obtiene el resultado satisfactorio de

$$g = v$$

La velocidad del grupo de ondas de materia es justamente igual a la velocidad de la partícula que gobiernan, y el postulado de de Broglie es internamente consistente. Utilizando expresiones relativistas para E y p, para evaluar dE/dp, se obtiene la misma conclusión.

A continuación se derivarán las relaciones de incertidumbre, combinando las relaciones de de Broglie-Einstein, $p = h/\lambda$ y $E = h\nu$, con las propiedades de grupos de ondas. Primero se considerará un caso límite simple. Sea λ la longitud de onda de de Broglie asociada con una partícula. Se puede representar una longitud de onda determinada (monocromática) en términos de una onda senoidal simple que se extiende a todos los valores de x, es decir, una onda sin modulación infinitamente larga como

$$\Psi = A \operatorname{sen} 2\pi(\kappa x - \nu t)$$

o

$$\Psi = A \cos 2\pi(\kappa x - \nu t)$$

Si la longitud de onda tiene definitivamente el valor λ, no existe la incertidumbre $\Delta\lambda$ y el impulso asociado a la partícula $p = h/\lambda$ será también definitivo, de modo que $\Delta p_x = 0$. En una onda tal, la amplitud tiene un valor constante A en todas partes; es la misma para toda la gama infinita de x. Por lo tanto, la probabilidad de encontrar a la partícula, que según Born está relacionada con la amplitud de la onda, no está concentrada en un intervalo particular de valores de x. En otras palabras, la posición de la partícula se desconoce totalmente. La partícula puede estar en cualquier parte, de modo que $\Delta x = \infty$. Análogamente se puede afirmar que como $E = h\nu$ y como la frecuencia tiene un valor definido, entonces $\Delta E = 0$. Pero, para asegurar que la amplitud de la onda es perfectamente constante en el tiempo, ésta debe ser observada durante un tiempo infinito, de modo que $\Delta t = \infty$. Para este caso sencillo, se satisface que $\Delta p_x \Delta x \geq h/2$, y $\Delta E \Delta t \geq h/2$, en los límites $\Delta p_x = 0$, $\Delta x = \infty$, y $\Delta E = 0$, $\Delta t = \infty$.

Para poder tener una onda cuya amplitud varíe con x o t, se deben superponer varias ondas nonocromáticas con longitudes de onda o frecuencias diferentes. Cuando se superponen dos ondas de ese tipo, se obtiene el fenómeno familiar de pulsos, como se vio anteriormeote en esta sección, con la amplitud modulada de modo regular a todo tiempo y en todo el espacio. Si se desea construir una onda con extensión determinada en el espacio (un solo grupo con un principio y un fin específico), entonces se deberán superponer ondas senoidales que posean un espectro continuo de longitudes de onda en el intervalo $\Delta\lambda$. La amplitud de un grupo tal, será cero fuera de la región de tamaño Δx

Para visualizar mejor lo anterior, considérese un número finito de ondas senoidales con longitudes de onda λ, o longitudes de onda recíproca κ, ligeramente diferentes. En la figura 3-9 se muestran siete ondas senoidales componentes $\Psi'_\kappa = A_\kappa \cos 2\pi(\kappa x - \nu t)$, a tiempo $t = 0$. Las longitudes de onda recíprocas, $\kappa = 1/\lambda$ toman sólo los valores enteros desde $\kappa = 9$ hasta $\kappa = 15$. La amplitud de cada onda está dada pnr A_κ con $A_{12} = 1$, $A_{13} = A_{11} = 1/2$, $A_{14} = A_{10} = 1/3$, y $A_{15} = A_{19} = 1/4$, tal como se muestra en la figura. El esquema se centra en $x = 0$, donde todas las ondas están en fase (por esta razón se utilizan cosenos), pero ocurren desfasamientos entre ellas, a medida que se desplazan de esa posición, en cualquier dirección. Consecuentemente, su suma $\Psi' = \Psi'_9 + \cdots + \Psi'_{15}$ oscila con máxima amplitud en $x = 0$, pero las oscilaciones decaen al aumentar o disminuir x a medida que las relaciones de fase entre las ondas componentes se complican. Se se adopta la convención usual de definir la extensión del grupo entre el punto de máxima amplitud y aquel cuya amplitud es la mitad de la máxima, entonces la superposición contiene un grupo cuya extensión espacial Δx puede obtenerse de la figura y es ligeramente mayor que $2\pi/12$. Con una convención análoga, el intervalo de longitudes de onda recíprocas que comprenden el grupo de ondas, $\Delta\kappa$ tiene el valor 1. Obsérvese que el valor aproximado del producto $\Delta x \Delta\kappa$ es $\pi/6 \simeq 1/2$. En el extremo derecho de la figura se indica la presencia de un grupo auxiliar, de la misma forma que el grupo central. Estos grupos auxiliares se forman en intervalos uniformemente espaciados a lo largo de las partes positiva y negativa del eje x y ocurren debido a que si se tiene un número finito de ondas componentes, existen puntos en el eje, separados de $x = 0$ por distancias iguales a un número entero diferente de longitudes de onda, para cada componente. En estos puntos las componentes están de nuevo en fase y el grupo se repite. Si en un intervalo fijo de longitudes de onda recíproca $\Delta\kappa$ se duplica el número de ondas componentes, la anchura del grupo central permanecerá esencialmente invariante, pero las distancias que lo separan de los grupos auxiliares se duplicará.

Si se combina un número infinito de ondas senoidales componentes, cada una con una longitud de onda recíproca infinitesimalmente diferente pero contenidas en el intervalo $\kappa = 9$ a $K = 15$, se obtiene un grupo central muy similar al que se muestra en la figura 3-9, pero los grupos auxiliares no estarán presentes. La razón es que en tales casos no existe una longitud en el eje x en la cual encajen un número entero exacto de longitudes de onda para cada una del infinito número de componentes. Todas las componentes están en fase, en y cerca de $x = 0$ y por lo tanto se combinan constructivamente para formar el grupo. Alejándose de este punto en cualquier dirección, las ondas componentes empiezan a desfasarse entre sí, porque sus longitudes de onda o longitudes de onda recíproca son diferentes. Más allá de ciertos puntos, las fases del número infinito de componentes se vuelven aleatorias, y por lo tanto, la suma de las ondas componentes será cero. Además, nunca más volverá a estar en fase. Por lo tanto, las componentes forma un grupo de extensión restringida Δx. Resulta obvio que entre mayor sea el intervalo de longitudes de onda recíproca $\Delta\kappa$ de donde se obtienen las componentes que forman el grupo, menor será la extensión del grupo Δx, la razón es simplemente que si las longitudes de onda cubren un intervalo más amplio, las fases se volverán aleatorias en una distancia muy corta. De hecho, Δx es justamente inversamente proporcional a $\Delta\kappa$. El valor exacto de la constante de proporcionalidad y la forma del grupo de ondas, dependen de las amplitudes relativas de las ondas componentes.

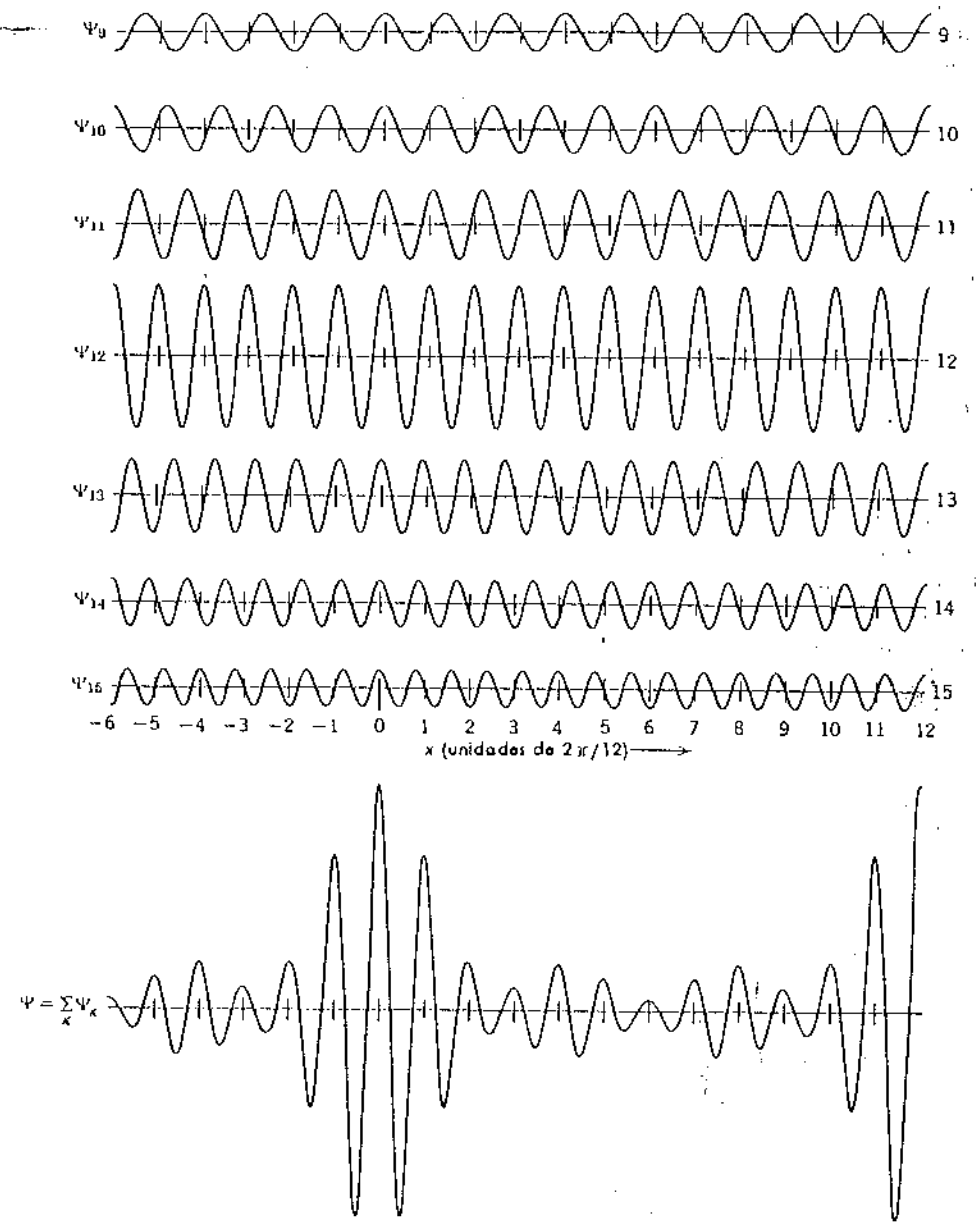


FIGURA 3-9

Se muestra, a $t = 0$, la superposición de siete ondas cosenoidales $\Psi_\kappa = A_\kappa \cos 2\pi(\kappa x - \nu t)$ con longitudes de onda recíprocas uniformemente espaciadas en el intervalo de $\kappa = 9$ a $\kappa = 15$. Sus amplitudes A_κ se maximizan en el valor $A_{12} = 1$ para la onda cuya κ está en el centro del intervalo y decaen simétricamente a los valores 1/2, 1/3, 1/4 para las otras ondas cuyas κ se aproximan a los extremos del intervalo. La suma $\Psi = \sum_\kappa \Psi_\kappa$ de estas ondas, consiste de un grupo centrado en $x = 0$. Definiendo Δx como la anchura entre la amplitud máxima y la mitad de la amplitud máxima de Ψ, y $\Delta\kappa$ se define como el intervalo de longitudes de onda recíprocas de las componentes de Ψ entre la amplitud máxima y la mitad de la amplitud máxima, se tiene $\Delta x \simeq 2\pi/12$, $\Delta\kappa \simeq 1$, y $\Delta x\, \Delta\kappa \simeq \pi/6 \simeq 1/2$.

Las matemáticas que se utilizan para llevar a cabo el procedimiento descrito, involucran la llamada integral de Fourier que es un poco más complicada, pero conducen a resultados numéricos muy similares a los obtenidos de la construcción de la figura 3-9. Es decir

$$\Delta x \Delta \kappa \geq 1/4\pi \tag{3-14}$$

El trabajo *óptimo* que pueda realizarse al formar un grupo de longitud Δx (amplitud media a la mitad de la altura) a partir de componentes cuyas longitudes de onda recíprocas cubren el intervalo $\Delta\kappa$ (mitad de la anchura a la mitad de la amplitud máxima) se obtiene $\Delta x = 1/4\pi\Delta\kappa$, o bien, $\Delta x \Delta\kappa = 1/4\pi$. Generalmente, de este producto se obtiene un valor algo mayor.

Un grupo de ondas que viaje en una región limitada del espacio pasará por algún punto de observación dado, en un tiempo limitado. Si el pulso o grupo de ondas tiene una duración t, entonces, necesariamente deberá componerse de ondas senoidales cuyas frecuencias cubran el intervalo $\Delta\nu$, donde

$$\Delta t \Delta \nu \geq 1/4\pi \tag{3-15}$$

Por lo tanto, la frecuencia del grupo tiene una dispersión $\Delta\nu$ si su duración es Δt, del mismo modo como existe una incertidumbre en su longitud de onda recíproca $\Delta\kappa$ si su anchura es Δx. La ecuación (3-15) también se obtiene de una integral de Fourier y con la ecuación (3-14), constituyen expresiones diferentes de la misma propiedad; sin embargo, para el estudiante será más familiar la relación frecuencia-tiempo o al menos algunas de sus implicaciones, como se muestra en el siguiente ejemplo.

Ejemplo 3-4. La señal de una estación de televisión consiste de pulsos cuya duración es $\Delta t \sim 10^{-6}$ seg. Explique por qué no es factible transmitir televisión en la banda AM.

El intervalo de frecuencias presentes en la señal, utilizando (3-15), es $\Delta\nu \sim 1/10^{-6}\,\text{seg} = 10^{6}\,\text{seg}^{-1} = 10^{6}$ Hz. Por lo tanto, la banda de transmisión en su totalidad ($\nu \simeq 0.5\times10^{6}$ Hz a $\nu \simeq 1.5\times10^{6}$ Hz) sería capaz de aceptar un solo "canal" de televisión. También se tendrían problemas serios con la construcción de transmisores y receptores con una banda fraccional tan amplia. A las frecuencias que se utilizan en transmisión de televisión ($\nu \simeq 10^{8}$ Hz) se pueden tener muchos canales en una región razonable del espectro, y sólo se requiere una *banda* nominal. ◀

Las ecuaciones (3-14) y (3-15) son *propiedades universales de todas las ondas* y se aplican a las ondas de materia, en combinación con las relaciones de de Broglie-Einstein, inmediatamente se obtienen las relaciones de incertidumbre de Heisenberg. Es decir, si en:

$$\Delta x \Delta \kappa = \Delta x \Delta (1/\lambda) \geq 1/4\pi$$

se sustituye $p = h/\lambda$ o $1/\lambda = p/h$,

$$\Delta x \Delta (p/h) = (1/h)\,\Delta x \Delta p \geq 1/4\pi$$

o

$$\Delta p \Delta x \geq h/2 \tag{3-16}$$

y si en

$$\Delta t \Delta \nu \geq 1/4\pi$$

se sustituye $E = h\nu$ o $\nu = E/h$, se obtiene

$$\Delta t \Delta (E/h) = (1/h)\,\Delta t \Delta E \geq 1/4\pi$$

o

$$\Delta E \Delta t \geq h/2 \tag{3-17}$$

Estos resultados concuerdan con las definiciones originales (3-5) y (3-6).

Resumiendo, se ha visto que en mediciones físicas, necesariamente se implica una interacción entre el observador y el sistema bajo observación. Para medidas de ese tipo, los entes disponibles son la materia y la radiación. Las relaciones $p = h/\lambda$ y $E = h\nu$, se aplican tanto a la radiación como a la materia, constituyendo así la expresión de la dualidad onda-partícula. Cuando estas relaciones se combinan con las propiedades universales a todas las ondas, se obtienen las relaciones de incertidumbre. Por lo tanto, el principio de incertidumbre es una consecuencia necesaria de esta dualidad, es decir, de las relaciones de de Broglie-Einstein; y el principio de incertidumbre en sí constituye la base de la tesis de Heisenberg-Bohr de que la probabilidad es fundamental para la física cuántica.

Ejemplo 3-5. Un átomo puede radiar en cualquier momento después de ser excitado. En un caso típico, se encuentra que la vida media de un átomo excitada promedio es de 10^{-8} seg. Es decir, durante este tiempo emite un fotón y sale de su estado de excitación.

(a) ¿Cuál es la incertidumbre $\Delta\nu$ mínima en la frecuencia del fotón?

De (3-15) se tiene

$$\Delta\nu\Delta t \geq 1/4\pi$$

o

$$\Delta\nu \geq 1/4\pi\Delta t$$

con $\Delta t = 10^{-8}$ seg, se obtiene $\Delta\nu \geq 8 \times 10^{6}\ \text{seg}^{-1}$.

(b) En átomos de sodio, la mayoría de los fotones corresponden a dos líneas espectrales de aproximadamente λ=5890 Å. ¿Cuál será la fracción de la amplitud $\Delta\nu/\nu$ correspondiente a cada línea?

Para $\lambda = 5890$ Å se obtiene $\nu = c/\lambda = 3 \times 10^{10}\ \text{cm-seg}^{-1}/5890\times10^{-8}\ \text{cm} = 5.1\times10^{14}\ \text{seg}^{-1}$. Por lo tanto $\Delta\nu/\nu = 8\times10^{6}\ \text{seg}^{-1}/5.1\times10^{14}\ \text{seg}^{-1} = 1.6 \times 10^{-8}$ o aproximadamente 2 partes en 100 millones.

Esta es la llamada *amplitud natural* de la línea espectral. En la práctica, la línea es mucho más gruesa por el ensanchamiento Doppler y el ensanchamiento por presión debidos al movimiento y a las colisiones de los átomos en la fuente.

(c) Calcule la incertidumbre en la energía ΔE del estado excitado del átomo.

La energía del estado excitado no se puede medir con precisión ya que sólo se tiene un tiempo finito para realizar la medida. Es decir, el átomo no permanece en el estado excitado un tiempo infinito, sino que decae a su estado de más baja energía y en ese proceso emite un fotón. De acuerdo con el principio de conservación de la energía, la dispersión en la energía del fotón debe ser igual a la dispersión en la energía del estado excitado del átomo. Si Δt es igual a la vida media del estado excitado, de (3-17) se tiene

$$\Delta E \geq \frac{h/4\pi}{\Delta t} = \frac{h}{4\pi\Delta t} = \frac{6.63 \times 10^{-34}\ \text{joule-seg}}{4\pi \times 10^{-8}\ \text{seg}}$$

$$= \frac{4.14 \times 10^{-15}\ \text{eV-seg}}{4\pi \times 10^{-8}\ \text{seg}} \simeq 3.3 \times 10^{-8}\ \text{eV}$$

Desde luego que lo anterior concuerda con lo obtenido en la parte (a), multiplicando la incertidumbre en la frecuencia del fotón $\Delta\nu$ por h para obtener $\Delta E = h\Delta\nu$.

La dispersión en energía de un estado excitado, por lo general se le llama *amplitud* del estado.

(d) De los resultados anteriores, determine la energía del estado excitado del átomo de sodio, hasta una precisión ΔE, con respecto a la energía de su estado más bajo, que emite un fotón cuya longitud de onda está centrada en 5890 Å.

Se tiene $\Delta\nu/\nu = h\Delta\nu/h\nu = \Delta E/E$. Por lo tanto $E = \Delta E/(\Delta\nu/\nu) = 3.3\times10^{-8}\ \text{eV}/1.6\times10^{-8} = 2.1$ eV, para lo cual han utilizado los cálculos de las partes (b) y (c). ◀

Ejemplo 3-6. Se hace una medida de la coordenada x de un electrón que pertenece a un ancho haz de electrones que se mueven en la dirección z, introduciendo en el haz una rendija angosta de amplitud Δy.

Demuestre que como resultado, se introduce una incertidumbre Δp_y en la componente y del impulso del electrón, de modo tal que $\Delta p_y \Delta y \geq \hbar/2$, como lo requiere el principio de incertidumbre. Efectúelo, considerando la difracción de la onda asociada con el electrón.

La onda, al propagarse por el aparato que se muestra en la figura 3-10, será difractada por la rendija. El ángulo θ correspondiente al primer mínimo del patrón de difracción de "una sola rendija", que se bosqueja en la figura, está dado por $\operatorname{sen}\theta = \lambda/\Delta y$. (Este es otro ejemplo de la relación general $\theta \simeq \lambda/a$ entre el ángulo de difracción, longitud de onda y dimensión característica del aparato de difracción). Como la propagación de la onda gobierna el movimiento de la partícula asociada, el patrón de difracción también da las probabilidades relativas de que los electrones lleguen a los diferentes puntos de la placa fotográfica. Por lo tanto, el electrón que pase por la rendija será deflectado por un ángulo que esté entre $-\theta$ y $+\theta$. A pesar de que su componente y del impulso había sido determinada con gran precisión igual a cero, antes de que pasara por la rendija (porque se sabía muy poco acerca de su posición y), después de atravesar la rendija, donde se realiza la medida de la posición y el impulso y puede ser cualquiera en el intervalo de $-p_y$ a $+p_y$, donde $\operatorname{sen}\theta = p_y/p$. De modo que la incertidumbre en el impulso y del electrón se debe a la medida de la posición y por la difracción de la onda electrónica.

La incertidumbre es:

$$\Delta p_y \simeq p_y = p \operatorname{sen}\theta = p\lambda/\Delta y$$

Usando la relación de de Broglie $p = h/\lambda$ para conectar el impulso de la partícula con la longitud de onda de la onda, se obtiene:

$$\Delta p_y = h/\Delta y$$

o

$$\Delta p_y \Delta y = h$$

Este resultado concuerda con el límite impuesto por el principio de incertidumbre. La difracción, que implica ondas, y el principio de incertidumbre que se refiere a partículas, proveen modos alternativos aunque equivalentes de tratar éste y todos los problemas similares. ◀

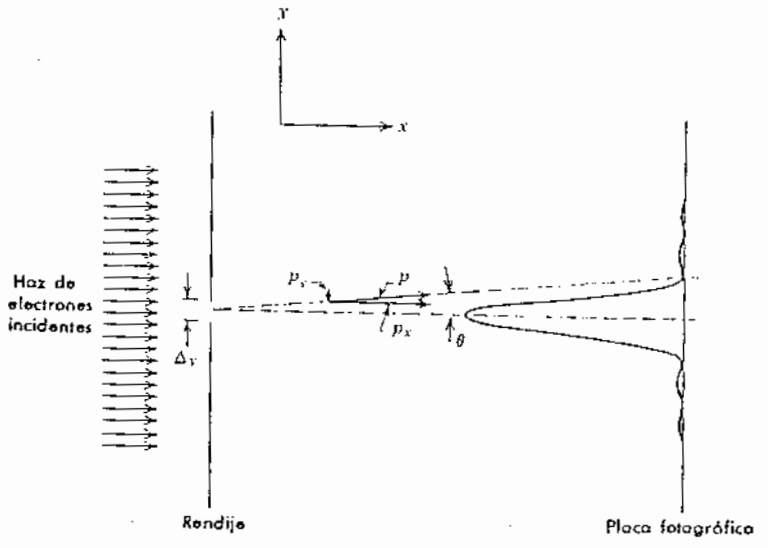


FIGURA 3-10

Medida de la coordenada y de un electrón en un haz paralelo y ancho, que se requiere que pase a través de una rendija. Se indica el patrón de difracción de la onda electrónica difractada, utilizando la línea que representa a la placa fotografiada, como un eje de la gráfica del patrón.

Obsérvese que en el ejemplo 3-6, se considera que la onda asociada con *un solo* electrón se difracta. La probabilidad de que el electrón pegue en algún punto de la placa fotográfica se determina por la intensidad de la onda electrónica. Si es sólo un electrón el que pasa por el aparato, podrá pegar en cualquier lugar del patrón de difracción, excepto en los puntos correspondientes a intensidad cern y seguramente pegará en algún punto cercano al máximo principal. Si son muchos los electrones que pasan por el aparato, cada una de sus ondas es difractada independientemente y los puntos sobre la placa fotográfica se distribuyen siguiendo el mismo patrón. El hecho de que los fenómenos de difracción impliquen interferencia entre partes diferentes de la onda correspondiente a una partícula y no a interferencia entre partes diferentes de ondas correspondientes a diferentes partículas, fue demostrado experimentalmente por primera vez, por G.I. Taylor para el caso de fotones y ondas de luz. Obtuvo un patrón de difracción utilizando luz de tan baja intensidad que los fotones pasaban por el aparato de difracción uno a uno, requiriendo para ello tiempos de exposición largos. Entonces, subiendo la intensidad a niveles normales de modo que hubiese muchos fotones en el aparato a cualquier tiempo, obtuvo el mismo patrón de difracción. Este mismo experimento ha sido esencialmente repetido posteriormente para electrnnes y otras partículas.

3.5 Algunas consecuencias del principio de incertidumbre

El principio de incertidumbre permite entender por qué es posible la naturalezadual (onda-partícula) de la radiación y la materia. Si se trata de determinar experimentalmente, si la radiación es una onda o una partícula, se encuentra que un experimento que fuerce a la radiación a revelar su carácter ondulatorio, suprime fuertemente su carácter corpuscular. Si el experimento se modifica para que revele el carácter corpuscular, se suprimen las características ondulatorias. En una misma situación experimental, no es posible confrontar los aspectos ondulatorio y corpuscular. Tanto la radiación como la materia semejan monedas que pueden desplegar una u otra cara a voluntad, pero nunca ambas simultáneamente. Desde luego que esta es la esencia del principio de complementaridad de Bohr: los conceptos de onda y partícula en lugar de contradecirse, se complementan.

Considérese el experimento de Young, de interferencia de luz por dos rendijas. Desde el punto de vista ondulatorio, la onda original es dividida en dos frentes de onda coherentes por las rendijas, y el traslape de estos frentes de onda es lo que produce las franjas de interferencia en la pantalla, que son características de todo fenómeno ondulatorio. Ahora supóngase que se sustituye la pantalla por una superficie fotoeléctrica. Si se mide el lugar de la superficie donde son emitidos los fotoelectrones, se obtiene un patrón de intensidades correspondiente al de dos rendijas presentándose así, aparentemente, aspectos ondulatorios. Sin embargo, si se mide las distribuciones de energía y tiempo de los fotoelectrones emitidos, se obtienen evidencias que demuestran que la radiación consiste de fotones, de modo que los aspectos corpusculares parecerán estar presentes. Si después se piensa en la radiación como fotones cuyo movimiento está gobernado por las propiedades de propagación de ciertas ondas asociadas (de Broglie), se llega a otra paradoja aparente. Cada fotón deberá pasar por una u otra rendija; si este es el caso, ¿cómo es posible que su movimiento más allá de las rendijas sea influenciado por la interacción de sus ondas asociadas con una rendija por la cual no pasó?

El error en el razonamiento de la paradoja se encuentra en la afirmación de que el fotón debe pasar por una u otra rendija. ¿Cómo puede determinarse realmente, en un experimento, si un fotón que se detecta en la pantalla, pasó por la rendija de arriba o la de abajo? Para poderlo hacer sería necesario colocar un detector en cada rendija, pero el detector que interacciona con el fotón en la rendija lo saca de la trayectoria que de otra forma seguiría. Por el principio de incertidumbre, se puede demostrar que un detector que posea suficiente resolución espacial como para poder determinar por cuál rendija pasa el fotón, perturba tanto su impulso que destruye el patrón de interferencia por dos rendijas. En otras palabras, si de algún modo se

demuestra que el fotón pasó por una rendija o la otra, entonces ya no se tendrá un patrón de interferencia. Si se desea observar el patrón de interferencia, debe evitarse perturbar a los fotones y no tratar de observarlos como partículas a lo largo de sus trayectorias hacia la pantalla. Se puede observar el comportamiento ondulatorio o corpuscular de la radiación; sin embargo, el principio de incertidumbre impide que se observen ambos al mismo tiempo, de modo que este comportamiento dual en realidad no es contradictorio. Esto también es aplicable al comportamiento onda-partícula de la materia.

El principio de incertidumbre también hace evidente la necesidad de que los sistemas cuánticos se expresen en términos de probabilidad. En la mecánica clásica, si se conoce exactamente la posición y el impulso de cada partícula, en algún instante, en un sistema aislado, entonces se puede predecir el comportamiento exacto de las partículas del sistema, en todo tiempo futuro; sin embargo en mecánica cuántica, el principio de incertidumbre muestra que lo anterior es imposible de hacer en sistemas que implican distancias e impulsos pequeños, ya que es imposible saber con la precisión requerida, las posiciones e impulsos instantáneos de las partículas. Consecuentemente, sólo se podrán predecir comportamientos *probables* de estas partículas.

Ejemplo 3-7. Considere una partícula microscópica que se mueve libremente a lo largo del eje x. Supanga que en el instante $t = 0$, se mide la posición de la partícula con una incertidumbre Δx_0. Calcule la incertidumbre en la posición de la partícula a un tiempo t posterior.

La incertidumbre en el impulso de la partícula, a $t = 0$, es al menos

$$\Delta p_x = \hbar/2\Delta x_0$$

Por lo tanto, la velocidad de la partícula, en ese instante, será incierta por lo menos

$$\Delta v_x = \Delta p_x/m = \hbar/2m\Delta x_0$$

y la distancia x que viajó la partícula en el tiempo t no podrá conocerse con mayor precisión que

$$\Delta x = t\Delta v_x = \hbar t/2m\Delta x_0$$

Si por una medida a $t = 0$, se localiza a la partícula dentro de un intervalo Δx_0, entonces, si se realiza una medida de su posición al tiempo t, la partícula podría encontrarse en un intervalo al menos tan grande como Δx.

Obsérvese que Δx es inversamente proporcional a Δx_0, de modo que entre *más* localizada esté la partícula en el instante inicial, será *menos* lo que se pueda saber de su posición final. Además, la incertidumbre Δx aumenta linealmente con el tiempo t. Esto corresponde a una dispersión, a medida que transcurre el tiempo del grupo de ondas asociado con el movimiento de la partícula. ◀

3.6 Filosofía de la teoría cuántica

A pesar de que existe un acuerdo entre todos los físicos, de que la teoría cuántica funciona en el sentido de que predice resultados que concuerdan en forma excelente con el experimento, existe una controversia cada vez mayor en cuanto a sus fundamentos filosóficos. Neils Bohr ha sido el principal arquitecto de la interpretación actual de la mecánica cuántica, conocida como la *interpretación de Copenhague*. Su enfoque ha sido apoyado por la gran mayoría de los físicos teóricos de hoy. Sin embargo, un grupo numeroso de físicos, no todos de acuerdo entre sí, han cuestionado la interpretación de Copenhague. El principal crítico de esta interpretación fue Albert Einstein. Los debates Einstein-Bohr constituyen una parte fascinante de la historia de la física. Bohr consideraba que había podido enfrentar todos los retos que inventó Einstein a manera de experimentos de pizarrón que tenían por objeto refutar el principio de incertidumbre.

Finalmente, Einstein aceptó que la teoría poseía consistencia lógica y su concordancia con los hechos experimentales, pero permaneció firme hasta el final, sin convencerse de que ello representara la realidad física última y dijo: "Dios no juega a los dados con el universo", refiriéndose a que la mecánica cuántica abandonaba los eventos individuales y la causalidad estricta, en favor de una interpretación fundamentalmente estadística.

Heiseberg ha afirmado el punto de vista comúnmente aceptado en forma concreta: "No hemos supuesto que la teoría cuántica, en forma opuesta a la teoría clásica, sea esencialmente una teoría estadística, en el sentido de que de los datos exactos sólo se puedan obtener conclusiones estadísticas.... En la formulación de la ley causal, a saber, si sabemos el presente exactamente, podemos predecir el futuro, no es la conclusión sino la premisa la que es falsa. No *podemos* saber, como cuestión de principio, el presente en todos sus detalles".

Entre los críticos del punto de vista de Bohr-Heisenberg de una indeterminación fundamental en la física, está Louis de Broglie. En un prólogo a un libro de David Bohm, joven colega de Einstein cuyos intentos de una teoría nueva revivieron el interés por reexaminar las bases filosóficas de la teoría cuántica, de Broglie escribe: "Razonablemente podemos aceptar que la actitud adoptada, por más de 30 años, por físicos teóricos cuánticos es, al menos en apariencia, la contraparte exacta de la información del mundo atómico que nos ha dado la experimentación. Al nivel de investigación que se tiene actualmente en microfísica, es cierto que los métodos de medición, no nos permiten determinar simultáneamente todas las cantidades que serían necesarias para obtener una visión de los corpúsculos de tipo clásico (esto se puede deducir del principio de incertidumbre de Heisenberg) y que las perturbaciones introducidas por la medición, que son imposibles de eliminar, en general nos impiden predecir el resultado que producirán y permiten sólo predicciones estadísticas. Por lo tanto, la construcción de fórmulas puramente probabilísticas, que todos los teóricos usan actualmente, fue completamente justificada. Sin embargo, la mayoría de ellos, muchas veces bajo la influencia de ideas preconcebidas derivadas de la doctrina positivista, han pensado que podrían ir más allá y afirmar que el carácter incierto e incompleto del conocimiento que la experimentación, en su estado actual, nos provee de lo que realmente ocurre en microfísica, es el resultado de una indeterminación real de los estados físicos y su evolución. Tal extrapolación de ninguna manera parece ser justificada. Es posible que en el futuro, examinando a un nivel más profundo la realidad física, seamos capaces de interpretar las leyes de probabilidad y física cuántica como los resultados estadísticos del desarrollo de variables completamente determinadas que actualmente se encuentran ocultas a nosotros. Puede ser que los poderosos medios que empezamos a usar para romper la estructura nuclear y para que aparezcan nuevas partículas, nos den algún día un conocimiento directo que actualmente no tenemos a este nivel más profundo. Tratar de impedir todo intento de ir más allá del punto de vista actual sobre la física cuántica, podría ser muy peligroso para el progreso de la ciencia y además sería contrario a las lecciones que podemos aprender de la historia de la ciencia. En efecto, esto nos enseña que el estado actual de nuestro conocimiento es siempre provisional y que deben existir, más allá de lo que actualmente se conoce, inmensas regiones nuevas que descubrir". (*Tomado de Causality and Chance in Modern Physics* por David Bohm, © 1957 D. Bohm; reimpreso con permiso de D. Van Nostrand Co.).

El estudiante deberá notar aquí la aceptación de que la mecánica cuántica es correcta al nivel atómico y nuclear. La búsqueda de un nivel más profundo, donde la mecánica cuántica podría ser superada, es motivada mucho más por la objeción a su indeterminismo filosófico, que por otras consideraciones. De acuerdo con Einstein, "La creencia en un mundo externo independiente del sujeto que lo percibe es la base de toda la ciencia natural". Sin embargo, la mecánica cuántica considera las interacciones entre objeto y observador como la realidad última. Utiliza el lenguaje de las relaciones físicas y procesos en lugar de el de las cualidades y propiedades físicas. Rechaza, por insignificante e inútil, el concepto de que detrás del universo de nuestra percepción se encuentre oculto un mundo objetivo gobernado por la causalidad; en cambio, se

confina a la descripción de las relaciones entre percepciones. De todas maneras, existen muchos que rehusan dejar de atribuir propiedades objetivas a las partículas elementales, por ejemplo, y trabajar con el conocimiento subjetivo acerca de ellas, lo cual motiva su investigación por una teoría nueva. De acuerdo con de Broglie, tal investigación es en el interés de la ciencia. Que lo anterior lleve a una nueva teoría que en algún campo actualmente inexplorado contradiga a la mecánica cuántica y altere también sus fundamentos filosóficos, nadie lo sabe.

PREGUNTAS

1. ¿Por qué la naturaleza ondulatoria de la materia no es más evidente en nuestras observaciones cotidianas?
2. ¿La longitud de onda de de Broglie solamente se aplica a "partículas elementales", tales como el electrón o el neutrón, o se aplica también a sistemas de materia complejos que poseen una estructura interna? Dé ejemplos.
3. Si en la fórmula de de Broglie, se hace $m \to \infty$, ¿se obtiene el resultado clásico para partículas macroscópicas?
4. ¿La longitud de onda de de Broglie puede ser menor que una dimensión lineal de la partícula? ¿Mayor? ¿Necesariamente existe una relación entre tales cantidades?
5. ¿La frecuencia de una onda de de Broglie se obtiene de E/h? ¿La velocidad es igual a $\lambda\nu$? ¿La velocidad es igual a c? Explique su respuesta.
6. ¿Se puede medir la frecuencia ν de una onda de de Broglie? Si la respuesta es afirmativa diga cómo.
7. ¿Cómo es que la difracción de electrones se puede utilizar para estudiar las propiedades de la superficie de un sólido?
8. ¿Cómo se explican los haces reflejados de modo regular en experimentos de difracción de electrones y átomos?
9. ¿Debe modificarse la fórmula de Bragg para electrones, para poder tomar en cuenta la refracción de ondas electrónicas en la superficie del cristal?
10. ¿Se obtiene una información diferente acerca de los cristales por medio de experimentos de difracción de electrones que por difracción de rayos X? ¿Dé experimentos de difracción de neutrones? Analícelo.
11. ¿Se podrían realizar estudios cristalográficos con protones? ¿con neutrones?
12. Analice la analogía: la óptica física es a la óptica geométrica como la mecánica ondulatoria es a la mecánica clásica.
13. ¿Es un electrón una partícula? ¿es una onda? Explique su respuesta.
14. ¿La longitud de onda de de Broglie asociada a una partícula del movimiento del sistema de referencia depende del observador? ¿Qué efectos tiene ésto, en la dualidad onda-partícula?
15. Dé ejemplos de cómo el proceso de medición perturba al sistema que se mide.
16. Demuestre la relación entre la naturaleza incontrolable del retroceso de Compton en el experimento del microscopio de rayos γ de Bohr y el hecho de que existen cuatro incógnitas y sólo tres ecuaciones de conservación en el efecto Compton.
17. Algunas veces, el principio de incertidumbre se enuncia en términos de cantidades angulares como $\Delta L_\varphi \Delta\varphi \geq \hbar/2$ donde ΔL_φ es la incertidumbre en una componente de impulso *angular* y $\Delta\varphi$ es la