

# Cauchy-Schwarz y el principio de incertidumbre

Fernando Chamizo

30 de enero de 2019

## Diálogos del físico desafortunado y la matemática indolente

FD—¡El principio de incertidumbre es alucinante! En física cuántica sabemos probar matemáticamente —no te quejarás— que es imposible medir simultáneamente la posición y el momento.

MI—No vamos a discutir como siempre acerca de lo que significa “probar”. Creo que los dos estaremos de acuerdo en que algo es una prueba dependiendo de lo que te creas.

FD—Por supuesto, tú no podrías probar el teorema de Pitágoras si no te crees los axiomas de Euclides. De hecho hay geometrías alternativas, como bien sabes, en las que conviene cambiarlos y entonces el teorema no es cierto. Por ejemplo, los triángulos curvos trazados sobre una superficie esférica como la Tierra, útiles desde hace mucho para la navegación, no cumplen el teorema.

MI—Imagina que uno de los axiomas que supones es que el cuadrado de la hipotenusa menos el cuadrado del cateto menor da el cuadrado del cateto mayor. Demostrarías el teorema muy fácilmente.

FD—Eso sería trampa, porque creerse tal cosa es tan difícil como creerse el teorema.

MI—Donde quiero llegar es que lo que me parece alucinante del principio de incertidumbre es la interpretación de posiciones y momentos que dieron nuestros primeros padres cuánticos: Bohr, Born, Heisenberg y otros incluyendo a Schrödinger y von Neumann que metieron muchas matemáticas en ello.

FD—Realmente Born, Heisenberg y Jordan representaron posiciones y momentos con matrices infinitas, una representación que pronto quedó obsoleta.

MI—Sí y su prueba era complicada y un poco dudosa. Lo que afirmo es que una vez que te crees la interpretación de los textos actuales básicos, la prueba matemática del principio de incertidumbre es prácticamente la desigualdad de Cauchy-Schwarz.

FD—Siempre tengo la impresión de que no solo estás amargada sino que quieres amargar a todo el mundo. Seguro que el principio de incertidumbre será menos bonito después de que me expliques todo esto.

MI—Me temo que te estás pasando a la zona oscura, ya empieces a pensar que menos matemáticas equivale a menos bonito.

FD—¡Ufff, eres una lianta! A veces me parece que los matemáticos creéis en dos mundos irreconciliables, el mundo real imperfecto donde ocurren desastres naturales y reina el caos y el mundo matemático perfecto, casi el de las ideas platónicas, mientras que los físicos teóricos pensamos que solo hay un mundo y está regido por leyes escritas en lenguaje matemático. ¿No es eso más positivo?

## 1. Cauchy-Schwarz a lo fácil

Comencemos viendo una prueba de la desigualdad de Cauchy-Schwarz distinta de la que habrás visto en clase, una prueba algebraica. Un momento, ¿no era la de clase ya algebraica? Sí, pero esta va a ser más algebraica en un sentido pedestre de la palabra porque solo involucra polinomios. En el caso de  $\mathbb{R}^n$  se reduce a la *identidad de Lagrange*

$$(1.1) \quad \sum_i x_i^2 \sum_j y_j^2 = \left( \sum_i x_i y_i \right)^2 + \frac{1}{2} \sum_i \sum_j (x_i y_j - x_j y_i)^2$$

donde  $i$  y  $j$  se mueven, digamos, entre 1 y  $n$ . En notación vectorial, para  $\vec{x}, \vec{y} \in \mathbb{R}^n$  esta identidad nos dice  $\|\vec{x}\|^2 \|\vec{y}\|^2 = (\vec{x} \cdot \vec{y})^2 + \text{algo positivo}$  que implica la desigualdad de Cauchy-Schwarz para el producto escalar de  $\mathbb{R}^n$ . Un producto escalar en un espacio vectorial de dimensión finita sobre  $\mathbb{R}$  siempre se reduce al usual con un cambio de coordenadas que pase a una base ortonormal. Por tanto a partir de (1.1) se tiene la desigualdad en espacios euclídeos reales cualesquiera de dimensión finita:

$$(1.2) \quad \|\vec{x}\| \|\vec{y}\| \geq |\langle \vec{x}, \vec{y} \rangle|.$$

Una prueba de (1.1) se sigue simplemente desarrollando el último cuadrado para obtener que el doble sumatorio es

$$(1.3) \quad \sum_i x_i^2 \sum_j y_j^2 + \sum_i y_i^2 \sum_j x_j^2 - 2 \sum_i x_i y_i \sum_j x_j y_j = 2 \sum_i x_i^2 \cdot \sum_i y_i^2 - 2 \left( \sum_i x_i y_i \right)^2.$$

Hay también una variante compleja de la identidad de Lagrange:

$$(1.4) \quad \sum_i |x_i|^2 \sum_j |y_j|^2 = \left| \sum_i \bar{x}_i y_i \right|^2 + \frac{1}{2} \sum_i \sum_j |x_i y_j - x_j y_i|^2$$

donde la barra indica el conjugado. La prueba de (1.4) es similar a la de (1.1) desarrollando el cuadrado  $|x_i y_j - x_j y_i|^2$  como  $x_i y_j - x_j y_i$  por su conjugado. Con ello tenemos (1.2) también para espacios vectoriales sobre  $\mathbb{C}$ .

En  $\mathbb{R}^3$  con el producto escalar usual, lo que dice (1.1) es

$$(1.5) \quad \|\vec{x}\|^2 \|\vec{y}\|^2 = |\vec{x} \cdot \vec{y}|^2 + \|\vec{x} \times \vec{y}\|^2$$

donde  $\vec{x} \times \vec{y}$  es el producto vectorial. A fin de cuentas, lo que está detrás de esta igualdad es  $\cos^2 + \sin^2 = 1$ . Si tienes el superpoder de visión en  $n$  dimensiones, geoméricamente lo que dice (1.1) es algo similar a (1.5) reemplazando  $\|\vec{x} \times \vec{y}\|$  por el área del paralelogramo que determinan  $\vec{x}$  e  $\vec{y}$ .

## 2. Una desigualdad de incertidumbre

A través de (1.1) nos percatamos de que las expresiones  $x_i y_j - x_j y_i$  son las que alejan (1.2) de ser una igualdad en  $\mathbb{R}^n$ . Si disponemos de una cota inferior para la suma de sus cuadrados, tendremos también una cota inferior para el producto de las normas. Con fórmulas:

$$(2.1) \quad \sum_i \sum_j (x_i y_j - x_j y_i)^2 \geq \lambda \quad \Rightarrow \quad \|\vec{x}\|^2 \|\vec{y}\|^2 \geq \frac{1}{2} \lambda.$$

Geoméricamente, en  $\mathbb{R}^3$  significa que el producto de las longitudes de dos vectores no puede estar nunca por debajo del modulo del producto vectorial, que es el área. Si  $x_i y_j - x_j y_i = 0$  para todo  $i, j \in \{1, 2, 3\}$ , el producto vectorial se vuelve conmutativo, y nulo, lo cual ocurre cuando los vectores son proporcionales. En ese caso no hay restricción para el producto de las longitudes.

Lo que vamos a ver es una variante de esta idea que involucra aplicaciones lineales (matrices) y la conclusión será que a no ser que dos aplicaciones conmuten, tendremos una cota inferior para el producto de ciertas normas. En la aplicación más famosa esas normas son las “precisiones” con las que se pueden medir posición y momento en mecánica cuántica y la existencia de una barrera inferior se traduce en que una gran precisión (norma pequeña) en una de estas magnitudes implica en una gran imprecisión (norma grande) en la otra.

Supongamos que trabajamos en un espacio vectorial  $V$  sobre  $\mathbb{C}$  y consideremos dos aplicaciones lineales hermíticas (autoadjuntas)  $A, B : V \rightarrow V$ . Es decir, identificando  $A$  y  $B$  con sus matrices en una base ortonormal, son matrices iguales a sus complejas conjugadas traspuestas. Si es la primera vez que oyes hablar de matrices hermíticas quizá te convenga pensar que estamos en  $\mathbb{C}^n$  con la base canónica y el producto escalar usual.

Dado  $\vec{v} \in V$ , definimos  $\vec{x} = A\vec{v}$ ,  $\vec{y} = B\vec{v}$ . Usando que son hermíticas,  $\langle \vec{x}, \vec{y} \rangle = \langle \vec{v}, AB\vec{v} \rangle$ , lo cual se puede escribir como

$$(2.2) \quad 2\langle \vec{x}, \vec{y} \rangle = \langle \vec{v}, \{A, B\}\vec{v} \rangle - i\langle \vec{v}, i[A, B]\vec{v} \rangle$$

con  $i = \sqrt{-1}$  donde  $\{A, B\} = AB + BA$  y  $[A, B] = AB - BA$ , unas notaciones que emplearás hasta la saciedad en cursos superiores de física. Es fácil ver que  $\{A, B\}$  e  $i[A, B]$  son también hermíticas. Ahora bien, si  $C$  es hermítica  $\langle \vec{v}, C\vec{v} \rangle \in \mathbb{R}$  porque  $\langle \vec{v}, C\vec{v} \rangle = \langle C\vec{v}, \vec{v} \rangle = \langle \vec{v}, C\vec{v} \rangle$ . Por tanto, cada uno de los productos escalares del segundo miembro de (2.2) son reales y se tiene tomando módulos al cuadrado

$$(2.3) \quad 4|\langle \vec{x}, \vec{y} \rangle|^2 \geq |\langle \vec{v}, [A, B]\vec{v} \rangle|^2.$$

Aplicando la desigualdad de Cauchy-Schwarz, el primer miembro es menor que  $4\|A\vec{v}\|^2\|B\vec{v}\|^2$  y se concluye

$$(2.4) \quad \|A\vec{v}\|\|B\vec{v}\| \geq \frac{1}{2}|\langle \vec{v}, [A, B]\vec{v} \rangle|.$$

De este modo, si  $A$  y  $B$  no conmutan tenemos una cota inferior no trivial para el producto de las normas que constituye el primer miembro. Como  $[A, B]$  es invariante al cambiar  $A \mapsto A - \mu I$  con  $I$  la identidad y  $\mu \in \mathbb{R}$  (con ello la aplicación sigue siendo autoadjunta), y lo mismo con  $B$ , sin esfuerzo adicional, (2.4) se generaliza a

$$(2.5) \quad \|A\vec{v} - \mu_1\vec{v}\|\|B\vec{v} - \mu_2\vec{v}\| \geq \frac{1}{2}|\langle \vec{v}, [A, B]\vec{v} \rangle| \quad \text{para cualesquiera } \mu_1, \mu_2 \in \mathbb{R}.$$

Si repasas la prueba de (2.5) verás que es fácil, poco más que Cauchy-Schwarz. Sí, puedes afirmar que es rara pero los pocos pasos de que consta se siguen sin dificultad. Pues bien, (2.5) incluye el principio de incertidumbre habitual y otros muchos. Lo difícil no es creerse esta desigualdad sino que posiciones y momentos tengan algo que ver con operadores hermíticos.

### 3. Teniendo fe en el diccionario cuántico

Los estados de los sistemas físicos se representan por funciones que toman valores complejos, llamadas *funciones de onda* muchas veces denotadas como  $\Psi$ . En situaciones simples, la  $\Psi$  que representa una partícula es una función de  $(x, y, z)$  y del tiempo  $t$ . Este último en lo que veremos se puede considerar fijo y por tanto nos olvidaremos de él. Para simplificar nos centraremos en el caso unidimensional y así tendremos  $\Psi = \Psi(x)$  en lugar de  $\Psi = \Psi(x, y, z)$  y  $\int$  significará  $\int_{-\infty}^{\infty}$  si no se dice lo contrario. No indicar los límites de integración deja la puerta abierta a que pongamos  $\mathbb{R}^3$  u otras cosas cuando seamos mayores. Las funciones de onda interesantes físicamente son aquellas que cumplen  $\int |\Psi|^2 = 1$  y así lo supondremos.

Hasta ahora, en los cursos de física el estado una partícula en un instante fijo venía dado por un punto (posición) y una velocidad (o mejor un momento). ¿Qué sentido tiene representarlo ahora mediante una función? No es que lo diga yo ni el telediario, es que lo dijo Schrödinger. Y esto es solo el principio. En esta sección enciende tu varita de incienso, relájate y si te crees el diccionario psicodélico que traduce conceptos físicos básicos a matemáticas superiores, te crearás en un santiamén el principio de incertidumbre. Cuando te pongas rebelde, pasa a la siguiente sección.

Se llaman *operador posición* y *operador momento* a los operadores  $A$  y  $B$  que actúan sobre las funciones de onda de la siguiente forma:

$$(3.1) \quad A : \Psi \mapsto x\Psi \quad \text{y} \quad B : \Psi \mapsto -i\hbar \frac{d}{dx} \Psi$$

donde  $\hbar \approx 1.0546 \cdot 10^{-34} \text{ J} \cdot \text{s}$  es la *constante de Planck normalizada* e  $i$  es el  $i = \sqrt{-1}$  de toda la vida. En física se suele escribir  $\hat{x}$  y  $\hat{p}$  en lugar de  $A$  y  $B$ . Así el operador posición aplicado a  $e^{-\pi x^2/2}$  da  $x e^{-\pi x^2/2}$  y el operador momento sobre la misma función de ondas da  $\pi i \hbar e^{-\pi x^2/2}$  ¡Toma ya! ¡Ahora posición y momento son operadores que pasan unas funciones a otras! Dijimos que íbamos a relajarnos, así que sigamos con un poco de matemáticas. Estos operadores son aplicaciones lineales cuando metemos las funciones de ondas en un espacio vectorial de funciones porque  $A(\Psi_1 + \Psi_2) = A(\Psi_1) + A(\Psi_2)$ ,  $A(\lambda\Psi) = \lambda A(\Psi)$  para  $\lambda \in \mathbb{C}$  y lo mismo con  $B$ . Los espacios de funciones tienen habitualmente dimensión infinita por lo cual no es muy buena idea intentar representar  $A$  y  $B$  con matrices<sup>1</sup>.

Sí, ya sé que no te lo quitas de la cabeza. ¿Cómo convencer al que pone los carteles en la carretera de distancias y límites de velocidad que lo que está midiendo son operadores? El truco está en que, si te fías de los padres fundadores de la mecánica cuántica, lo que uno ve en el mundo macroscópico es una especie de promedio de estos operadores para la función de ondas del sistema considerado. Esta *posición media* y *momento medio* son

$$(3.2) \quad \bar{x} = \int x |\Psi|^2 \quad \text{y} \quad \bar{p} = -i\hbar \int \bar{\Psi} \frac{d}{dx} \Psi$$

donde  $\bar{\Psi}$  significa el conjugado de  $\Psi$  mientras que las barras encima de  $x$  y  $p$  son, por supuesto, notación, una forma común de referirse a promedios. En el mundo subatómico, las mediciones de

<sup>1</sup>En el futuro verás que las funciones de ondas que representan el espín son un caso amable de dimensión finita en el que aparecen unas matrices  $2 \times 2$  famosas, llamadas *matrices de Pauli*.

posiciones y momentos pueden dar cualquier cosa (demasiado tarde para pedir y el libro de reclamaciones a los padres fundadores) y  $\bar{x}$  y  $\bar{p}$  son realmente la media de esas mediciones. En esta situación de resultados incontrolados, se vuelve relevante la *desviación típica*: cuán lejos se separan normalmente las medidas de los promedios. Una desviación típica pequeña significa que el promedio es muy orientativo mientras que cuando es grande se tiene una gran dispersión en las mediciones. Pues bien, el cuadrado de la desviación típica, la *varianza*, para la posición y el momento viene dada por los formulones

$$(3.3) \quad \sigma_x^2 = \int (x - \bar{x})^2 |\Psi|^2 \quad y \quad \sigma_p^2 = \int \left| -i\hbar \frac{d}{dx} \Psi - \bar{p} \Psi \right|^2.$$

Lo primero recuerda un poco a la fórmula de la varianza que quizá hayas visto en los cursos de estadística pero lo segundo no demasiado.

Fin de la psicodelia. Definamos el producto escalar<sup>2</sup>

$$(3.4) \quad \langle f, g \rangle = \int \bar{f} g$$

en el espacio vectorial sobre  $\mathbb{C}$  generado por todas las funciones de onda que usemos. La norma asociada es  $\|f\| = (\int |f|^2)^{1/2}$  y las fórmulas anteriores se escriben de manera bastante simétrica:

$$(3.5) \quad \bar{x} = \langle \Psi, A\Psi \rangle, \quad \bar{p} = \langle \Psi, B\Psi \rangle, \quad \sigma_x = \|(A - \bar{x} \text{Id})\Psi\|, \quad \sigma_p = \|(B - \bar{p} \text{Id})\Psi\|$$

con Id el operador identidad. No te dejes engañar por este aparente orden en el caos, los formulones (3.2) y (3.3) vienen de estas fórmulas y no al revés. Notemos además que

$$(3.6) \quad [A, B]\Psi = x \cdot (-i\hbar) \frac{d}{dx} \Psi - (-i\hbar) \frac{d}{dx} (x\Psi) = i\hbar \Psi.$$

Entonces (2.5) con  $\vec{v} = \Psi$  nos da<sup>3</sup>

$$(3.7) \quad \boxed{\sigma_x \sigma_p \geq \frac{\hbar}{2}}$$

que es el *principio de incertidumbre* y se deduce una limitación diminuta, pero relevante en rangos subatómicos, para la precisión al medir simultáneamente posición y momento. Según (2.5), la culpa de esta limitación recae en la falta de conmutatividad de  $A$  y  $B$  y por ello en  $[A, B]\Psi = i\hbar \Psi$ . En física se suele escribir esta última fórmula como

$$(3.8) \quad [\hat{x}, \hat{p}] = i\hbar$$

y se le llama *relación de indeterminación*.

<sup>2</sup>Con la notación del curso el producto escalar debe ser  $\int f \bar{g}$  en lugar de  $\int \bar{f} g$  pero exigir la linealidad en el segundo argumento está tan extendido en física que no me atrevo a cambiarlo.

<sup>3</sup>¿No habíamos hecho todo en dimensión finita? ¿Y si ahora no lo estamos? La respuesta rápida es que a un físico probablemente eso le importe poco por la idea intuitiva de que siempre se puede aproximar por dimensiones finitas y pasar al límite.

## 4. Nociones cuánticas mal contadas en un ratillo

El resumen de la sección anterior es que si te crees (3.5) entonces (3.7) es prácticamente directo a partir de (2.5). Lo que intenta dar esta sección es una idea de cómo concebir que las mediciones de posición y momento, ambos conceptos físicos básicos, tengan medias y desviaciones típicas dadas por una fórmulas tan raras. Para ello habrá un poco de historias de abuelete y más física que matemáticas.

A principios del siglo XX diversos experimentos relacionados sobre todo con la física atómica sugirieron que las partículas se comportaban como ondas. Por ejemplo, la famosa *hipótesis de de Broglie* asociaba a una partícula de momento  $p$  una longitud de onda  $\lambda = h/p$  donde  $h$  es la *constante de Planck*, esta vez sin normalizar, dada por  $h = 2\pi\hbar$ . Las ondas con esta longitud de onda son genéricamente  $e^{ipx/\hbar}$  donde la exponencial compleja es solo una manera matemáticamente cómoda de no tener que tratar separadamente los senos y cosenos, gracias a la conocida fórmula de Euler

$$(4.1) \quad e^{i\alpha} = \cos \alpha + i \operatorname{sen} \alpha.$$

También se puede interpretar como una forma de representar una fase.

Si uno atiende a la dependencia en el tiempo, en parte los experimentos y en parte la teoría sugerían que las ondas puras que corresponden a momento  $p$  y energía  $E$  son

$$(4.2) \quad \Psi_{p,E}(x,t) = e^{i(px-Et)/\hbar}.$$

La famosa *ecuación de Schrödinger*, uno de los postulados de la mecánica cuántica, afirma que la onda asociada a una partícula, la *función de ondas*  $\Psi$  del apartado anterior, es superposición de estas ondas puras  $\Psi_{p,E}$ . De este modo una partícula puede contener multitud de momentos y energías (que Born y Heisenberg indicaban con matrices infinitas). Cuando Schrödinger introdujo su ecuación, la mecánica cuántica era una mecánica ondulatoria que postulaba que las partículas eran realmente ondas<sup>4</sup> representadas por una función de ondas  $\Psi$  que, como antes suponemos siempre normalizada<sup>5</sup> con  $\int |\Psi|^2 = 1$ .

Curiosamente Schrödinger estaba confuso acerca del significado físico de  $\Psi$  e inicialmente creyó que para los electrones  $|\Psi|^2$  daba la densidad de carga. La interpretación actual, procedente de Born, es que fijado el tiempo,  $|\Psi|^2$  en cada punto es la densidad de probabilidad de detectar la partícula en dicho punto. A esto se une un extraño fenómeno llamado el *colapso de la función de ondas* que es motivo todavía de controversia porque combina partículas cuánticas con instrumentos de medición clásicos haciendo que el significado de medir sea bien diferente del habitual. Si quieres saber sobre esto, las palabras clave son *interpretación de Copenhague* y *problema de la medición*. En el tercer apartado de [Wei05], un premio Nobel te dice unas palabras sobre ello.

Aunque el colapso de la función de ondas es central para entender la mecánica cuántica, aquí nos restringiremos a la visión incompleta originaria puramente ondulatoria. Una primera pregunta

---

<sup>4</sup>De alguna manera la *segunda cuantización* que se utiliza en la actualidad al introducir campos, devuelve el protagonismo a las partículas [LB14, §2.1].

<sup>5</sup>Las ondas puras  $\Psi_{p,E}$  no están normalizadas, todavía peor  $\int |\Psi_{p,E}|^2 = \infty$ . Eso no debe inquietarnos porque son una idealización, no son la función de ondas de algo “real”. De hecho es natural que ocurra así porque indican algo así como una partícula que está en todos los sitios con igual probabilidad y si me dicen elige un número al azar en  $\mathbb{R}$  yo no sabría calcular la probabilidad de que esté en un conjunto dado. Si tú sabes, ve al médico de los infinitos.

natural es que si por ejemplo un electrón es una onda ¿por qué nos parece una partícula? La respuesta es que en las situaciones habituales es una onda muy concentrada. De la misma manera, nos parece que tiene un momento preciso porque en la superposición de las  $\Psi_{p,E}$  de (4.2) que da la función de ondas del electrón, hay solo un estrecho margen de valores de  $p$  que tiene coeficientes (amplitudes) sensiblemente grandes.

Por ejemplo, según la teoría, un electrón en el estado fundamental de un átomo de hidrógeno tiene como función de ondas

$$(4.3) \quad \Psi(r, t) = (\pi r_0^3)^{-1/2} e^{-r/r_0 - iE_1 t/\hbar} \quad \text{donde} \quad r_0 = 5.29 \cdot 10^{-11} m, \quad E_1 = 2.18 \cdot 10^{-18} J$$

y  $r$  es la distancia al núcleo. Esta  $\Psi$  se vuelve prácticamente nula por el decaimiento exponencial si  $r/r_0$  es grande, por tanto es imposible detectar la onda del electrón mucho más allá de  $r_0$ . Por cierto, no es nada fácil expresar esta  $\Psi$  en términos de las ondas puras (4.2) (con  $\vec{p} \cdot \vec{x}$  en lugar de  $px$  porque estamos en tres dimensiones) pero es posible sabiendo suficientes matemáticas y cuando se hace, se ve que la contribución de los momentos de norma mucho mayor que  $\hbar/r_0$  es despreciable. En la siguiente sección haremos algunos cálculos precisos.

¿Cómo averiguar la posición de una onda? En principio es imposible, así como  $x$  está en todas partes de la recta real pero todos seríamos capaces de decir dónde está aproximadamente una ola o una onda lo suficientemente concentrada, como en el caso del electrón anterior. Siguiendo la interpretación dada anteriormente,  $|\Psi(x)|^2$  es exactamente la densidad de probabilidad de detectar la partícula en  $x$  y si nos obligan a especificar con un solo número las infinitas posiciones de una onda más o menos concentrada, lo natural es elegir la *posición media*

$$(4.4) \quad \bar{x} = \int x |\Psi(x)|^2 dx.$$

Así que la primera fórmula de (3.2), y por tanto la primera de (3.5), no es tan rara una vez que uno ha creído a Born. Lo mismo se aplica a la primera fórmula de (3.3) que no es más que la usual de la varianza para la función de densidad  $f = |\Psi|^2$ .

Veamos que es medianamente comprensible o al menos explicable que llamemos “posición” al operador  $A$  de (3.1). Para ello pensemos que un pulso, un onda abrupta muy estrecha tiene una posición bastante bien definida. Troceando una onda en rodajas verticales siempre la podremos descomponer en una suma de pulsos muy finos. El operador posición simplemente lo que hace es multiplicar cada pulso por su posición, que está bien definida en el límite de anchuras infinitesimales.

Uno de los postulados de la mecánica cuántica es que los *observables* físicos son aplicaciones lineales autoadjuntas (hermíticas). Aunque no lo parezca, esto es un artificio teórico que refleja el hecho de que exista una descomposición de cualquier función de ondas de forma que el observable adquiera un valor de verdad, un número real, en cada trozo de la descomposición sobre el que actúa por multiplicación.

El operador momento es más complicado porque la descomposición de una función de ondas  $\Psi$  en términos de los trozos  $\Psi_{p,E}$  que tienen momento bien definido es matemáticamente bastante más difícil que hacer rodajas verticales. Para simplificar, supondremos que la energía  $E$  es constante, igual que ya lo habíamos hecho con el tiempo. Si  $\Psi$  es superposición de muchas  $\Psi_{p_j,E}$  y, como en las

rodajas anteriores, los  $p_j$  pueden estar arbitrariamente juntos, es mejor tomar como modelo que  $\Psi$  es una integral en vez de una suma<sup>6</sup>:

$$(4.5) \quad \Psi(x) = \int a(p)\Psi_{p,E}(x) dp.$$

Con la filosofía anterior, como  $\Psi_{p,E}$  tiene momento  $p$ , el operador momento debe ser:

$$(4.6) \quad B : \Psi \mapsto \int pa(p)\Psi_{p,E}dp.$$

Esto no tiene mucha pinta de parecerse a (3.1), entre otras cosas porque no tenemos ni idea de qué es  $a(p)$ . La explicación es mucho más sencilla de lo que parece, solo hay que saber derivar exponenciales:

$$(4.7) \quad \int pa(p)\Psi_{p,E}(x)dp = -i\hbar \frac{d}{dx} \int a(p)\Psi_{p,E}(x)dp = -i\hbar \frac{d}{dx} \Psi.$$

Ahora hay un hecho nada obvio y es que (4.5) implica  $h \int |a|^2 = 1$ . Esto se deduce tomando  $f_1 = f_2 = \Psi$  y  $a_1 = a_2 = a$  en la fórmula matemática alucinante

$$(4.8) \quad f_1(x) = \int a_1(p)\Psi_{p,E}(x)dp, \quad f_2(x) = \int a_2(p)\Psi_{p,E}(x)dp \quad \Rightarrow \quad h \int \bar{a}_1 a_2 = \int \bar{f}_1 f_2.$$

La palabra mágica es *trasformada de Fourier* y solo la menciono para intrigar porque no habrá explicaciones sobre ella. Como  $h \int |a|^2 = 1$ , se puede considerar que  $h|a|^2$  es una densidad de probabilidad y, en analogía con la interpretación de Born, (4.5) induce a pensar que  $h|a(p)|^2$  es la densidad de probabilidad de medir momento  $p$  en el sistema físico representado por  $\Psi$ . Con ello, el momento medio y su varianza están dados por

$$(4.9) \quad \bar{p} = h \int p|a|^2 \quad y \quad \sigma_p^2 = h \int (p - \bar{p})^2 |a|^2$$

en completa analogía con  $\bar{x}$  y  $\sigma_x^2$ . De nuevo esto se parece bien poco a sí mismo, esto es, a las fórmulas para  $\bar{p}$  y  $\sigma_p^2$  en (3.2) y (3.3). El cálculo que explica esta paradoja no es complicado pero sí hermético porque vuelve a usar (4.8). Veámoslo

Tomando  $a_2 = pa$  se tiene  $f_2 = -i\hbar \frac{d}{dx} \Psi$  por (4.7) y si además elegimos  $a_1 = a$ , de (4.8) se sigue  $-i\hbar \int \bar{\Psi} \frac{d}{dx} \Psi = h \int p|a|^2$ , que es lo que queríamos probar para  $\bar{p}$ . La igualdad entre las dos fórmulas para  $\sigma_p^2$  se sigue de la misma forma tomando esta vez  $a_1(p) = a_2(p) = (p - \bar{p})a(p)$  que por (4.7) corresponde a  $f_1 = f_2 = -i\hbar \frac{d}{dx} \Psi - \bar{p}\Psi$ .

## 5. Un ejemplito con un electrón

Para practicar y explorar un poco el caso tridimensional, vamos a calcular posiciones y momentos para el electrón del átomo de hidrógeno correspondiente a (4.3). En este caso hay posiciones y

<sup>6</sup>En los ejemplos que estudiarás en cursos superiores los habrá tanto con sumas como con integrales e incluso combinando ambas. Cuando solo hay sumas, momentos y energías solo puede tomar unos valores concretos, por ejemplo relacionados con los enteros, y se dice que están cuantizados.



momentos correspondientes a cada una de las tres coordenadas. Por ejemplo

$$(5.1) \quad \bar{y} = \int_{\mathbb{R}^3} y |\Psi|^2 \quad \text{o} \quad \bar{p}_z = -i\hbar \int_{\mathbb{R}^3} \bar{\Psi} \frac{\partial}{\partial z} \Psi.$$

Al sustituir la  $\Psi$  de (4.3) con  $r = \sqrt{x^2 + y^2 + z^2}$  es fácil ver que  $\bar{x}$ ,  $\bar{y}$ ,  $\bar{z}$ ,  $\bar{p}_x$ ,  $\bar{p}_y$  y  $\bar{p}_z$  son todos nulos ya que se integra una función impar en una de las variables. Esto es lógico con la imagen clásica del electrón dando vueltas alrededor del núcleo, pues la posición media estará en el centro (el origen) y lo mismo con el momento. Las desviaciones típicas no son, por supuesto, nulas. Por ejemplo, teniendo en mente (3.3)

$$(5.2) \quad \sigma_z^2 = \int_{\mathbb{R}^3} z^2 |\Psi|^2 = \frac{1}{\pi r_0^3} \int_{\mathbb{R}^3} z^2 e^{-2\sqrt{x^2+y^2+z^2}/r_0} dx dy dz = r_0^2.$$

La última igualdad se sigue pasando a coordenadas esféricas o preguntando al profesor de Análisis II. La segunda fórmula de (3.3) para  $p_z$  da

$$(5.3) \quad \sigma_{p_z}^2 = \hbar^2 \int_{\mathbb{R}^3} \left| \frac{\partial \Psi}{\partial z} \right|^2 = \frac{\hbar^2}{\pi r_0^5} \int_{\mathbb{R}^3} \frac{z^2 e^{-2\sqrt{x^2+y^2+z^2}/r_0}}{x^2 + y^2 + z^2} dx dy dz = \frac{\hbar^2}{3r_0^2}.$$

La desigualdad (2.4) para los operadores  $A_z : \Psi \mapsto z\Psi$  y  $B_z : \Psi \mapsto -i\hbar \frac{\partial}{\partial z} \Psi$  asegura que en cualquier ejemplo  $\sigma_z \sigma_{p_z} \geq \hbar/2$ , ya que  $[A_z, B_z] = i\hbar$ . En nuestro ejemplo con el electrón del átomo de hidrógeno en el estado fundamental hemos obtenido  $\sigma_z \sigma_{p_z} = \hbar/\sqrt{3}$  que está alrededor de un 15% por encima del límite. Dicho sea de paso, hay una prueba matemática de que solo se alcanza la igualdad en (3.7) cuando  $|\Psi|^2$  es de la forma  $C e^{-C^2 \pi (x-K)^2}$  con  $C > 0$  y  $K$  constantes. Es trivial comprobar que  $A_y : \Psi \mapsto y\Psi$  y  $B_z$  conmutan,  $[A_y, B_z] = 0$ . En este caso no hay ningún principio de incertidumbre que impida la existencia de sistemas físicos en los que la posición en la dirección  $y$  y el momento en la dirección  $z$  tengan precisión arbitrariamente pequeña. Por ejemplo para  $|\Psi|^2 = e^{-\pi(x^2+y^2/\epsilon+\epsilon z^2)}$  se cumple  $\sigma_y^2 = \epsilon^{3/2}/(2\pi)$  y  $\sigma_{p_z}^2 = \hbar^2 \pi \epsilon^{1/2}/2$  que son arbitrariamente pequeños cuando  $\epsilon \rightarrow 0^+$ .

Por la simetría se tiene  $\sigma_x = \sigma_y = \sigma_z = r_0$  y  $\sigma_{p_x} = \sigma_{p_y} = \sigma_{p_z} = \hbar/(r_0\sqrt{3})$ . Eso explica que, como se indicó antes, sea muy inusual detectar el electrón mucho más lejos de  $r_0$  del núcleo o con momento mucho mayor que  $\hbar/r_0$ . Debido al decaimiento exponencial de  $|\Psi|^2$  las probabilidades de que ocurra esto se vuelven muy pequeñas en cuanto que “lejos” sea “moderadamente lejos”. Por ejemplo, la probabilidad de encontrar al electrón a una distancia mayor que  $10r_0$  del núcleo es

$$(5.4) \quad \int_{r>10r_0} |\Psi|^2 = \frac{1}{\pi r_0^3} \int_{x^2+y^2+z^2>10r_0} e^{-2\sqrt{x^2+y^2+z^2}/r_0} dx dy dz = 221e^{-20}.$$

Esta cantidad es menor que media millonésima.

## 6. Para saber más

En [SR05] y [Bag17] hay mucha información histórica acerca de cómo se gestaron las ideas fundacionales de la física cuántica. También son muy ilustrativas las primeras páginas de [GP78]. Un resumen rápido está en [Wey50]. Es increíble que ese libro se escribiera en 1928 cuando todavía se estaban asimilando los conceptos básicos.

Originalmente la relación de indeterminación era una identidad entre productos de matrices infinitas. A pesar de que esta interpretación tuvo una historia muy corta, es instructivo (aunque nada sencillo) intentar seguir estos primeros pasos en que posiciones y momentos se convirtieron en operadores. Una explicación moderna detallada está en [FP09] y hay traducciones de algunos de los artículos originales en [vdW67].

## Referencias

- [Bag17] J. Baggott. *La historia del cuanto: una historia en 40 momentos*. Ediciones de Intervención Cultural, 2017.
- [FP09] W. A. Fedak and J. J. Prentis. The 1925 Born and Jordan paper “On quantum mechanics”. *Amer. J. Phys.*, 77(2):128–139, 2009.
- [GP78] A. Galindo and P. Pascual. *Mecánica cuántica*. Alhambra, Madrid, 1978.
- [LB14] T. Lancaster and S. J. Blundell. *Quantum Field Theory for the Gifted Amateur*. Oxford University Press, New York, 2014.
- [SR05] J. M. Sánchez-Ron. *Historia de la física cuántica. vol. 1: El periodo fundacional (1860-1926)*. Drakontos. Crítica, D.L.2001, Barcelona, 2005.
- [vdW67] B. L. van der Waerden. *Sources of quantum mechanics*. Dover Publications, Inc., New York, 1967. Classics of Science, Volume V.
- [Wei05] S. Weinberg. Einstein’s mistakes. *Physics Today*, 58(11):31–35, 2005.
- [Wey50] H. Weyl. *The theory of groups and quantum mechanics*. Dover Publications, Inc., New York, 1950. Translated from the second (revised) German edition by H. P. Robertson, Reprint of the 1931 English translation.